Aalto University

LULEÅ UNIVERSITY OF TECHNOLOGY

Rian van Gijlswijk

A feasibility study on ground-based localization for Mars exploration

School of Electrical Engineering

Department of Electrical Engineering and Automation

Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Technology

Espoo, Thursday 20th August, 2015

Instructor: Heikki Vanhamäki
University of Oulu
Astronomy and Space Physics unit

Supervisors: Professor Esa Kallio
Aalto University
School of Electrical Engineering

Dr. Thomas Gustafsson
Luleå University of Technology

Abstract

Aalto University
School of Electrical Engineering Abstract of the Master's Thesis

Author:	Rian van Gijlswijk
Title of the thesis:	A feasibility study on ground-based localization for Mars exploration
Date:	Thursday 20th August, 2015 **Number of pages:** 112
Department:	Department of Electrical Engineering and Automation
Programme:	Master's Degree Programme in Space Science and Technology
Professorship:	Automation Technology (AS-84)
Supervisors:	Professor Esa Kallio (Aalto)
	Dr. Thomas Gustafsson (LTU)
Instructor:	Heikki Vanhamäki

The goal of this work is to assert the feasibility of the use of a ground-based network, consisting of beacons, to aid in the provision of localization for mobile stations on Mars. The dayside ionosphere can be utilized to refract and reflect radio signals to some receiver beyond the horizon. The propagation of radio signals is studied with the use of a simulation, which makes use of a ray-tracing method to track radio signals in the ionosphere.
Several scenarios are defined to classify the state of the ionosphere. A nominal scenario is provided, with ionospheric properties prevalent at Mars. Furthermore, scenarios are defined to include both a minimum and maximum solar activity, as well as dust storms and SEP events. The determination of the position of a mobile station can be done by exploiting the properties of the received signal. Three localization technologies are studied in detail in this work, being the *time of arrival* (TOA), *received signal strength* (RSS) and *angle of arrival* (AoA). The *Cramer-Rao lower bound* (CRLB) is used to determine the lower bound of the precision of a position estimate, produced by any of these technologies. A comparison of the obtained precisions shows that it is possible to obtain a precision of less than $13m$ for any scenario when using TOA technology, and $6m$ when using AoA technology. Furthermore, a coverage of at least 92 % is achieved in all but the most extreme scenarios. The coverage increases with decreasing frequency. The best results are obtained with a frequency of $2.5MHz$. Moreover, it is shown that the best precision is obtained by placing beacons either in a rectangular or triangular grid on Mars, with beacons spaced 10° apart. By doing so, a $4m$ precision can even be obtained under the nominal scenario for a frequency of $2.5MHz$ and $4.5MHz$ with any technology. Ultimately, a ground-based beacon system has proven to be a sound solution to providing positioning estimates on the Martian surface.

Keywords: Mars, ionosphere, localization

Preface

This research is conducted at the school of Electrical Engineering at the Aalto University, in Finland.

I would like to express my gratitude to my supervisor, professor Esa Kallio. He supported me through the research from the moment of inception. His enthusiasm and great knowledge of the ionosphere were what made this research happen in the first place. Moreover, I would like to thank my advisor Heikki Vanhamäki, whose knowledge of the ionosphere and meticulous inspection of my work is highly appreciated. Next, I highly value the help and support I received from José Vallet, who explained me all the statistics I needed to come to a conclusion. His eye for precision is highly appreciated. Last, I would like to thank my girlfriend and family for their support during this research. Their enthusiasm and interest kept me motivated.

Ultimately, I believe that Mars is a prime target for both autonomous and human exploration. The interest in this planet has a long heritage, and even as much as a hundred years ago scientists, like Percival Lowell, believed that advanced live existed on the planet. Whether there is any life at all remains to be seen, but this question has certainly fueled and inspired many people. Recent scientific missions have shown that Mars is far more than a barren planet. Dried riverbeds and organic material depositions indicate that Mars is a place packed with mysteries yet to be discovered, and therefore a target worthy of exploration.

It is but a matter of time before humans set foot on the red planet to explore investigate, and possibly colonize. Any infrastructure that would support them in this endeavor is greatly appreciated. It is my intent to provide both humans and robots with the infrastructure needed to explore while they know where they are.

August, 2015

Rian van Gijlswijk

Contents

List of Tables

List of Figures

List of Acronyms

AoA	Angle of Arrival
BC	Beacon
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
CRLB	Cramer-Rao Lower Bound
DME	Distance Measuring Equipment
DR	DataRate
ED&L	Entry, Descent and Landing
EUV	Solar Extreme Ultraviolet
GDOP	Geometric Dilution of Precision
GPS	Global Positioning System
HDOP	Horizontal Dilution of Precision
HF	High-Frequency radio waves (1 to $10MHz$)
HPBW	Half-Power Beam Width
LOS	Line of Sight
LS	Least-Squares
MARSIS	Mars Advanced Radar for Subsurface and Ionosphere Sounding radar
MC&N	Mars Communication & Navigation
MCECS	Mars-Centered Ecliptic Coordinate System
MGS	Mars Global Surveyor
MS	Mobile Station
MSL	Mars Science Laboratory
MUF	Maximum Usable Frequency
NDB	Non-Directional Beacon
NLOS	Non-Line of Sight
PDF	Probability Density Function
PDOP	Position Dilution of Precision
RSS	Received Signal Strength
SEP	Solar Energetic Particle
SNR	Signal-to-Noise ratio
Sol	The duration of a day on Mars
SZA	Solar Zenith Angle
TDOA	Time Difference of Arrival
TDOP	Time Dilution of Precision
TECP	Thermal and Electrical Conductivity Probe
TOA	Time of Arrival
UHF	Ultra High Frequency

List of Symbols

A	Absorption
α	Azimuth angle
B	Channel bandwidth
β	Latitude
Bf	Blackout factor
c	Speed of light $2.997 \cdot 10^{8}\ [ms^{-1}]$
Δ	Elevation angle
dB	Decibel
d_n	Distance between beacon n and a mobile station
DR	Datarate
d_s	Skip distance
e	Electron charge $1.602 \cdot 10^{-19}\ [C]$
ϵ	Permittivity of free space $8.854 \cdot 10^{-12}[F/m]$
ϵ'	Real part of the permittivity
ϵ''	Imaginary part of the permittivity
f	Ray frequency
f_p	Plasma frequency
γ_a	Angle of arrival (elevation angle)
H_n	Neutral scale height
h_{max}	Electron peak density
k	Boltzman constant: $1.38 \cdot 10^{-23}[J/K]$
κ	Absorption per unit length
ξ	Index of absorption
λ	Longitude
m_e	Electron mass $9.109 \cdot 10^{-31}\ [kg]$
μ	Index of refraction
μ_{Bf}	Mean blackout factor
$\mu_{\hat{p}}$	Mean estimated precision
N	Substitute refractive index
n_c	Complex index of refraction
n_{CO2}	Number density of CO_2
$n_{CO2,s}$	Number density of CO_2 at surface
N_e	Electron number density
P_r	Power received $[dB]$
P_t	Power transmitted $[dB]$
ϕ	Launch elevation
ϕ_c	Iris angle
$\bar{r_d}$	Mean radius of a dust particle
$\boldsymbol{r_i}$	Incident ray
$\boldsymbol{r_r}$	Reflected ray
$\boldsymbol{r_t}$	Refracted ray
τ	Propagation time
θ_t	Refracted ray angle w.r.t. normal
θ_i	Incident ray angle w.r.t. normal
θ_r	Reflected ray angle w.r.t. normal
θ_t	Refracted ray angle w.r.t. normal
T_s	System temperature
ν	Collision frequency
ω	Ray angular frequency
$w(t)$	A window function of parameter t
Ω_c	Gyrofrequency
ω_p	Plasma angular frequency
χ	Solar Zenith Angle

1 — Introduction

Mars, a seemingly barren planet, is replete with mysteries that have yet to be explored. In the present day, ever more missions are performed to study the red planet. However, these missions are still subject to a plethora of limitations in various technological aspects. As such, science is limited. In order to enhance support for exploration on Mars, the technological readiness level must be raised. Doing so not only broadens the scope of scientific research but perhaps one day also enables a manned mission.

One of these limitations is found in the field of navigation. The autonomy of a rover is to a significant extent determined by its ability to determine its position and navigate to a point of interest. Currently rovers are not able to autonomously navigate on the red planet but instead rely on directions from science teams on Earth to achieve their goal. The need for improved positioning information in all these missions is imperative. Moreover, the operational range of a rover is greatly increased if an accurate position information is available. This will increase the autonomy of a rover as it will be able to determine its path without intervention from Earth.

1.1 Problem description

Mars, having neither significant magnetic poles nor a satellite-based positioning constellation has limited possibilities for localization. Compasses do not work and a GPS fix is naturally not available. Currently localization is done by maintaining and updating a map with landmarks of the exploration area of a rover (Parker, 2013). The map is used for daily science planning and the position of the rover is redetermined while it waits for new instructions. This method is time consuming and involves the interference of a science team on Earth. Without this limitation the range of a rover will be greatly extended as the rover can transverse large distances without the need to wait for instructions.

1.2 Thesis objective

The objective of this thesis is to explore a method of providing positioning information to missions on the Martian surface. This method will be done using a beacon-based positioning network. The beacons use the ionosphere to reflect radio signals, so that beyond-the-horizon reception of a beacon signal becomes possible.

The Martian horizon is very close by and hence a rover or human-controlled vehicle would only be able to make line-of-sight contact with beacons closer than $3km$. This would greatly reduce the feasibility of ground-based beacons as one would need many to provide adequate coverage. However, the reflective properties of the ionosphere can be exploited to provide beyond-the-horizon reception of beacon signals. This method is already applied in the aviation and marine industry by the means of *non-directional beacons* (NDB). As such prior knowledge on the technology is readily available. This proposal will adapt the existing technology for the Martian environment so that positioning of rovers is also possible when the beacons are beyond the visible horizon from the rovers' point of view.

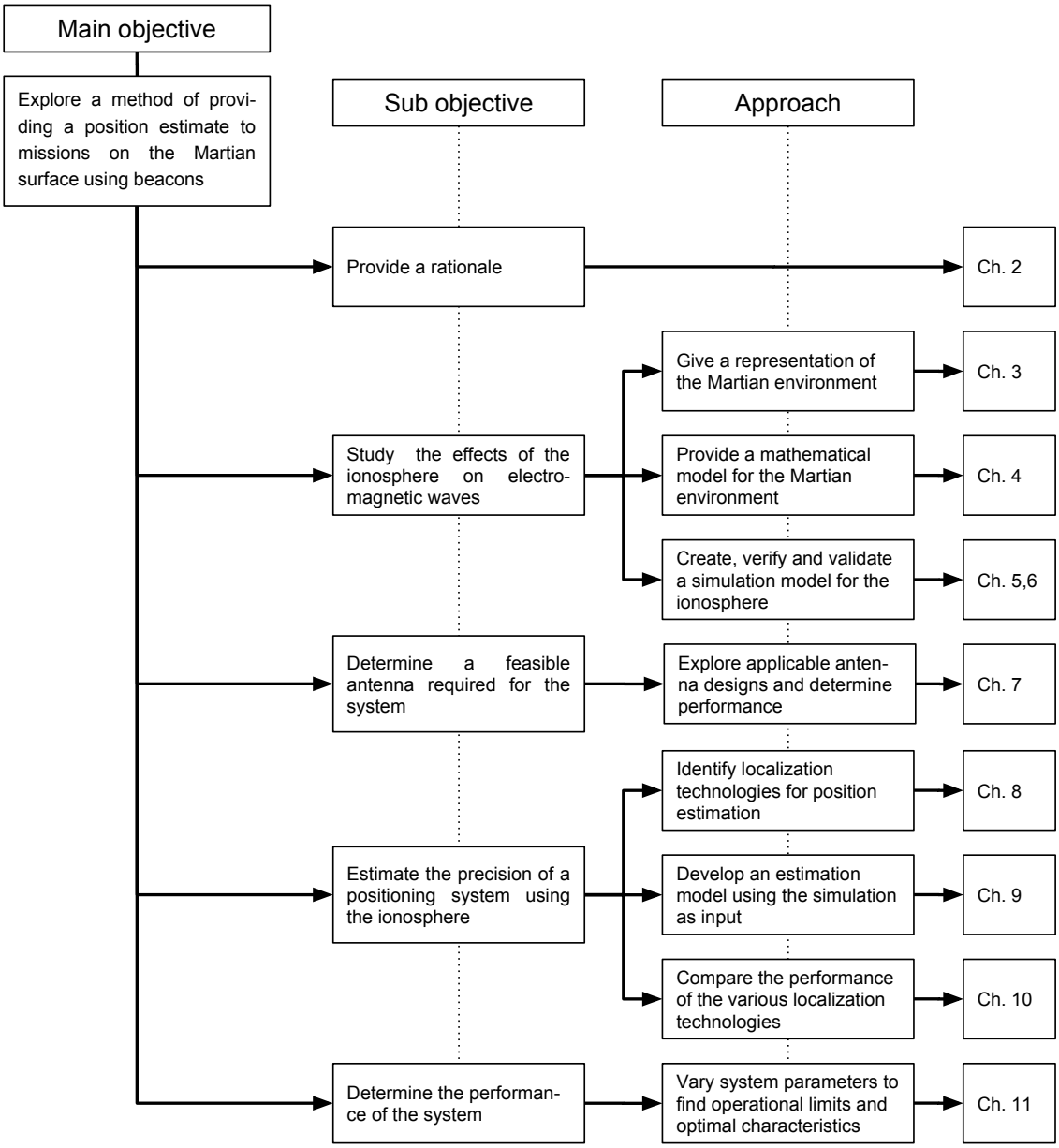


Figure 1.1: Report structure

1.3 Scope

The thesis covers three focus areas in order to complete the objective, which combined explore the feasibility of a beacon based localization network on Mars. These three areas are:

- Study the effects of the ionosphere
- Estimate the precision of localization using the ionosphere
- Determine the requirements for transmitters and receivers used in the system

It is imperative to study each three of these focus areas so that the performance of the system can be determined. The thesis will therefore conclude with a performance estima-

tion following from the studies of the focus areas.

1.4 Report structure

The report is structured as represented in the roadmap in figure 1.1. In this roadmap, the main objective is dissected into various sub-objectives or focus areas. First, the ionosphere is researched, so that a tool can be developed which simulates the ionosphere. This tool is subsequently validated and verified. Afterwards, the implications of antenna design are studied. In the third focus area, an approach is developed to estimate the precision of a positioning system which uses the ionosphere for radio propagation. Ultimately, the performance of such a positioning system is tested under various ionospheric conditions and system parameters, leading to the conclusion of this work.

2 — Thesis Motivation

Present-day Mars missions have a limited scientific yield because the mission hardware is only partially autonomous. Moreover, missions rely on external support when it comes to exploring the Martian terrain. There is an increased need for more precise navigation and position techniques, driven by the need for more sophisticated and bigger missions to Mars, eventually leading to human exploration. A viable approach to solve this issue is to study how the same problems have been solved back on Earth. Navigation and position techniques for marine and aviation applications on Earth have been meticulously studied for a long time now. On the other hand, little has been studied on the same problem on Mars.

This chapter provides the rationale as to what positioning system is most viable to use in a Martian environment. It does so by first identifying the limitations experienced by present-day Mars missions. Next, it explores the requirements that drive the need for positioning on Mars. Subsequently the related terrestrial position techniques and their application on Earth are analyzed. Also, a brief overview of current proposals for positioning systems on Mars is given. Finally, a conclusion is made as to which method best suits as a positioning system on Mars, based on a comparison from the existing and proposed methods described in this chapter. It is found that a ground-based system of beacons using the ionosphere lends the best perspective for providing a position estimate on Mars.

2.1 Methods used on current Mars missions

Position estimation is a cumbersome process for current Mars missions. Due to the long propagation delay between Earth and Mars, teleoperation of rovers cannot be done in real-time and thus becomes a challenge. Therefore, a certain level of autonomy is required. Present-day missions rely on a combination of dead reckoning, inertial navigation and support from scientists on Earth in order to determine the position of a rover (Parker, 2013). In the case of the rovers Spirit & Opportunity, as well as Curiosity, a stereo image pair taken by the onboard stereo camera can be used to determine the position of the rover with respect to a 3D landscape and correct for the cumulation of localization errors caused by odometry (Powell et al., 2006). However, this either requires extensive onboard processing or a significant amount of bandwidth to upload the image pairs to Earth for further analysis. In the latter case the rover stays idle and waits for new instructions while scientists on Earth use the stereo image pairs to work out its position and make a planning for the next Sol.

2.2 Drivers for positioning methods

The exploration of Mars will see an influx of more sophisticated science missions than ever before, driving the need for better navigation techniques. Whereas past and current missions rely on onboard systems and support from science teams back on Earth, future missions will require new levels of autonomy that cannot be met with the current techniques.

To determine what levels of autonomy are required, missions can be classified according to their positioning requirements. Mars missions can generally be divided into three types

(Edwards et al., 2001):

- Remote sensing orbiters
- Scout class missions
- Large entry class systems

The subsequent section will assess the needs for each of these three classes. Table 2.1 concludes with an overview of all needs categorized per class.

2.2.1 Remote sensing orbiters

Remote sensing orbiters are tasked with mapping the red planet. Their main technology drivers lie in the provision of remote sensing systems and an adequate communication link to relay the data (Edwards et al., 2001). Obviously, they do not need a position estimate, as the orbit dictates the position. Nevertheless, remote sensing orbiters might play a supporting role for ground-based navigation systems. While orbiters do not require a position estimate themselves, they can carry instruments which guide other spacecraft as they approach the planet or navigate down to the surface. Ever since the Mars Reconnaisance orbiter, such an instrument is provided on board orbiters which gathers doppler data used for navigation of other missions (Graf et al., 2005). This instrument can estimate the position of spacecraft as they approach Mars and aerobrake in its atmosphere by measuring the doppler shift of the approaching spacecraft. As a result, more precise data will be gathered of the orbit of this spacecraft after aerobraking. In fact, this method also helps with determining more accurately the final position of a lander as it is landing on the surface.

2.2.2 Scout class missions

Scout class missions are small landers, probes, or aerobots, which are mainly constrained by energy and communication budgets (Edwards et al., 2001). Small landers and probes don't require continuous localization but the availability of a positioning system eases the task of locating them on the surface.

Another upcoming type of proposed scout class missions are aerobots. These bots have been proposed to supplement human exploration so that a greater area for exploration can be covered (Zubrin, 1992). Furthermore, aerobots would support novel rover missions by exploring the terrain ahead of the rover such that more information about the terrain is available beforehand. Such an exploration method is proposed for the upcoming MSL2020 mission (Volpe, 2014). The need for position information for aerobots is more evident. The MSL2020 aerobot requires position information as it needs to know where the rover is. Similar cases can be made for independent aerobots as they require realtime position information to navigate to a target of interest. Naturally, they cannot loiter for an entire Sol waiting for navigation instructions from Earth.

2.2.3 Large entry class systems

Large entry class systems have the most profound need for position information. Various mission types can be gathered this class, each with their own requirements.

First, large rovers such as Curiosity are bound by their limited sensors and processing power. As such they need information from Earth in order to navigate through a terrain (Parker, 2013). This means of navigation is slow, hence limiting the scientific yield which

can be obtained from these missions. The provision of position information to these rovers could greatly speed up their operations.

Second, a sample-and-return type mission is a type of mission which has been in the planning stages for a long time. The goal of such a mission is to bring back regolith from Mars. Various proposals have been made, most of which require a lander and a rover which gathers regolith and brings them back to the lander. The lander contains an ascend stage which will be launched back to Earth, its payload bays filled with regolith. This requires the rover to be able to move back to the lander. Another type of proposal suggest providing a lander with a drill (Karcz et al., 2012). Such a mission is stationary, but must land right on top of scientifically interesting regolith. In this case, it is desirable for the lander to perform a pinpoint landing.

Last, human exploration of Mars requires a whole new level of positioning and navigation capabilities. Humans travelling to Mars will set up an outpost and explore from there using manned and unmanned rovers (Drake, 2009). The distance that manned rovers can travel from the outpost is, among other things, limited by the ability to navigate back to the outpost. In order to enhance this distance, it is imperative to provide accurate position and navigation methods which will work under any circumstances, including nighttime conditions and dust storms. Furthermore, the outpost approach will require pinpoint landing such that landers with supplies or new crew can land precisely close to the existing outpost in a safe manner.

Table 2.1: Positioning requirements for various mission types

Mission class	Mission type	Requirements
Remote sensing orbiters	Any	Provide support
Scout class missions	Aerobots Rover support	Position information
Large entry class systems	Large rovers Sample return mission Human exploration	Pinpoint landing All-weather position information

2.3 Terrestrial systems

Terrestrial positioning systems can be dissected into three categories: radio systems, celestial systems and mapping systems (Kayton and Fried, 1997). Radio systems, such as GPS , provide positioning information through a constellation of ground-based or space-based beacons. A receiver can deduce its position by measuring various signal parameters such as time-of-flight, received signal strength or a doppler shift. Celestial systems determine their position by studying the positions of known celestial objects relative to a planetary surface. Star trackers are an example of such systems. Finally, mapping systems infer their position based on images of known landmarks. These systems are used on current Mars missions, and come with several limitations as described in section 2.1.

In this section, the study of applicable terrestrial systems is limited to radio systems only. The goal of this research is to overcome the limitations caused by mapping systems. Furthermore, celestial systems have their main use in high-altitude or space-based applications. As a result, the focus is on radio systems. These systems can again be divided in three

categories (Kayton and Fried, 1997):

- Point source systems
- Hyperbolic systems
- Space-based systems

All of these systems provide positioning information on Earth. This section will compare the benefits and drawbacks of each system if applied in the Martian environment. This way, a decision can be made as to which system is the most viable when used on Mars.

2.3.1 Point source systems

Point source systems, such as non-directional beacons (NDBs), are beacons which provide position information by broadcasting an omnidirectional signal (Kayton and Fried, 1997). The signal relies on the surface conductivity and atmospheric refraction for its propagation. Receivers can tune to the signal and determine their angle with respect to the beacon by assessing the angle of incidence of the NDB signal upon the receiver.

Many NDBs are equipped with distance measurement equipment (DME). This allows the receiver to determine both the relative distance and the angle with respect to the beacon. Combined, these parameters give a rough estimate of the position of the receiver. When multiple NDBs are within reach of the receiver, triangulation can be performed to yield a position estimate with a higher accuracy.

The application of a comparable system on Mars poses a few challenges. First, Mars has a higher curvature than Earth due to its smaller radius, requiring radio signals to be refracted more. However, the atmospheric refractivity is two orders of magnitude lower than on Earth (Ho et al., 2002). The higher curvature also results in a shorter line-of-sight distance. Second, ground conductivity plays an important role in the attenuation of radio signals. On Earth, seawater provides the best propagation properties, having a conductivity of $2S/m$ to $6S/m$ (Laby, 2015). Sand and desert are known to have low conductivities. A low conductivity will cause a faster decay in the field strength of a radio signal (Millington, 1949). This in turn will incur a higher attenuation of that radio signal as the attenuation is a function of the field strength (ITU-R P.368-7, 1992). Compared to the Terrestrial environment, Martian soil is expected to be quite insulating, having an expected conductivity in the order of $10^{-9}S/m$ to $10^{-14}S/m$ (Greeley and Haberle, 1991). Comparable results were found by the TECP instrument on the Phoenix lander, which detected no conductivity at all (Zent et al., 2010). As such, the Martian regolith is a bad propagator of ground waves.

However, the Martian ionosphere is relatively stable during the daytime such that radio signals are refracted back to the surface. Previous research has suggested that this refractivity can be used to provide beyond-the-horizon radio signal propagation (Fry and Yowell, 1994; Melnik and Parrot, 1999). Furthermore, the Martian ionosphere is significantly less dense than Earths' ionosphere (Ho et al., 2001). Therefore, radio signals are less attenuated in the Martian ionosphere, as the attenuation is a function of the ionospheric density (Nielsen et al., 2007).

2.3.2 Hyperbolic systems

Hyperbolic systems provide some added benefit in a Martian environment as opposed to point source systems. They rely on timing or phase differences from multiple beacons to determine a position along a hyperbolic line (Kayton and Fried, 1997). Just as with a point

source system, one beacon in a hyperbolic system broadcasts a signal which is guided by the atmosphere, ionosphere and surface. Two beacons can yield a position solution along a hyperbolic line, three beacons providing a unique position fix. More than three beacons can further increase the accuracy of the fix. Therefore, the challenges of point source systems can be remedied, which would otherwise cause inaccuracies.

The challenge in hyperbolic systems lies in the trade-off between the accuracy and the number of beacons. A higher number of beacons will increase the accuracy as position information from more beacons is available. However, this will also increase total mission cost and complexity. As the position information is only available within reach of the beacons, one would have to tailor such a system to the missions at hand. Furthermore, a challenge arises with deploying the beacons themselves, as they need to be operational on the Martian surface. However, this challenge might already be resolved. The design studies of the Mars-MetNet mission have demonstrated a simple method of deploying meterological stations on the Martian surface. In this mission, the stations are deployed using a novel ED&L technology which enables a high payload mass and high mission robustness (Harri et al., 2006).

2.3.3 Space-based systems

The application of a space-based system, such as a GPS constellation, poses a few challenges. First, GPS satellites require very precise knowledge of their orbits. The knowledge of the satellite orbits around Mars might not be precise enough to achieve a reasonable accuracy. Furthermore, such a system needs highly accurate clocks in order to provide an accurate position estimate. Terrestrial GPS is based on time. Groundstations on Earth have access to what is regarded as *true time.* GPS satellites are corrected daily by these groundstations to correct for drifts in their timekeeping. As a result, the groundstations ensure that the time is constantly synchronized accross the entire system. Naturally, on Mars there are no groundstations available to provide timekeeping service. This will impact the accuracy of a GPS constellation on Mars, as the constellation will have to be self reliant in timekeeping.

2.4 Proposed methods

A *Mars Communication & Navigation* (MC&N) constellation, proposed by Bell et al. (2000), showed that it is possible to obtain position information with a sub-$10m$ resolution from a satellite constellation akin to the GPS constellation on Earth. This constellation would use only 6 satellites and is able to operate with a bare minimum of 4 satellites, albeit at the cost of degraded performance. A drawback of this system is that it would take on average $1.5hr$ to obtain a sub-$10m$ resolution. As the radio signals propagate the Martian ionosphere, they become distorted and thus ionospheric effects must be included. A study by Mendillo et al. (2004) has shown that the total electron content has a severe impact on the accuracy of a Martian GPS constellation. However, his impact can be mitigated by using dual-frequency transmission in order to cope with the distortion of the radio signals. Another drawback of the proposed constellation is that the time it takes to acquire a position fix with a certain accuracy depends on the latitude. Around the Martian equator, the lowest delays are obtained. However, near the poles it will take significantly longer to obtain a fix.

2.5 Selected method

A comparison is made between the various systems in order to judge what kind of positioning system would most suit an application on Mars. This section takes into account the existing and proposed systems for terrestrial positioning and Mars missions respectively and compares their advantages and disadvantages. From the comparison an optimal system is found, which will be studied in more detail in the subsequent chapters of this report.

2.5.1 Performance metrics

The benefit of a positioning system depends on a couple of performance metrics. These are in part comparable to the performance metrics of terrestrial systems as defined by Kayton and Fried (1997), but some extra parameters need to be taken into account considering the different environment of Mars. These extra parameters follow from the discussion in section 2.2 where the drivers for the need of positioning information on Mars were pointed out. This yields the following list of desired performance metrics.

Accuracy The accuracy is defined by the offset between the actual and estimated position and the time it takes to achieve this offset, measured as the *root mean square* RMS position error. A positioning system needs to balance the achievable RMS error and the delay in the acquisition of this error. Many internal and external parameters can influence the magnitude of the RMS error. First, the propagation effects through the Martian environment will cause phase delays, attenuations and other malevolent effects. Second, the geometrical arrangement of transmitters causes *geometrical dilution of precision* (GDOP) effects where the accuracy differs in each dimension. Last, errors in the instrumentation itself can cause additional errors in the position estimation.

The time delay plays an crucial role in mission benefit on a number of cases. First, the discussion of limitations of current missions in section 2.1 demonstrated the need for timely information as waiting for positioning data results in a lower scientific output. Second, human-rated missions will be tasked with the ability of navigating humans on the red planet safely. This means that a large time delay is not desired in the case of emergencies or other actions where time is of the essence. Third, pinpoint landing is a requirement as mentioned in section 2.2. During the landing, phase instantaneous position information is required for a spacecraft to perform any reasonably accurate landing. In this light, long time delays are out of the question.

Coverage The coverage of a positioning system is defined by the area where the signal-to-noise (S/N) ratio is above a minimum level which end users can acquire. The required coverage depends on the locations of these users, Mars missions in the case of this study. Missions are few and far apart, but not spread out over the planet. Hence, a positioning system providing global coverage is not necessary. It is desirable to have a positioning system which can provide positioning information to those areas that matter, while not spilling resources such as transmit power trying to achieve global coverage. However, positioning systems should not be committed to only one single scientific mission. It is desirable to have the ability to expand the positioning system once more missions require its services. The advent of colonization will arguably require a positioning system which can provide global coverage. However, this is considered a futuristic prospect.

Availability Availability is defined as the probability that position information is available. This depends on the uptime and thus the robustness of the system. Human rated missions will require an availability high enough to enable the safe navigation of humans on the surface. As such, it deems wise to meet the same requirements as posed upon terrestrial applications, which is an availability of 99.7% (Kayton and Fried, 1997). The uptime can be negatively influenced if mission hardware breaks down. Therefore, it is desirable that the mission hardware can be easily replaced or repaired if a failure occurs.

Ambiguity Ambiguity occurs when one set of positioning information yields more than one possible position. When this occurs, a position cannot be uniquely identified. Hence, the avoidance of ambiguities is a necessary condition.

Mission complexity In the case of Mars, deploying a positioning system induces extra complexity because the entire system needs to be transported to Mars. The extra complexity will increase the total mission cost. Furthermore, an increase in mission complexity tends to increase the total payload mass, which requires larger launch vehicles, further increasing mission cost and complexity. Hence, an optimal positioning system for application on Mars is simple and lightweight.

2.5.2 Trade-off

A simple comparison is made to compare two options of a positioning system which result from the discussion in this chapter. The first system is a network of ground-based beacons which provides positioning information similarly to a hyperbolic positioning system as was described in sections 2.3.1 and 2.3.2. The second system is a space-based system derived from the proposed method of a MC&N by Bell et al. (2000) as discussed in section 2.3.3 and 2.4.

Assumptions Some assumptions have to be made in order to qualitively trade-off the two proposed systems. Therefore, it is assumed that the unit cost for ground-based beacons is lower than that of the space-based satellites. The viability in this assumption stems from the fact that more support systems are required for a satellite as opposed to a beacon. Satellites require subsystems for station keeping and orbit maintenance, thermal control and protection against the harsh environment of space. Beacons have no need for station keeping and orbit maintenance. They solely require a subsystem to protect against the Martian environment, which is considerably less harsh than space itself. It is true that ground-based beacons require an ED&L subsystem in order for them to deploy on the surface, but space-based systems will require a form of aerobreaking or propulsive breaking as well in order to obtain a stable orbit around Mars.

Discussion The benefits of both ground-based and space-based systems become apparent in different performance metrics. An overview hereof is given in table 2.2. A few specific drawbacks arise from the comparison given in the table.

First, in terms of accuracy it is highly questionable to what extend a space-based positioning system is able to provide pinpoint landing opportunities, considering the nominal

Table 2.2: A comparison of the benefits of two positioning systems on Mars

Metric	Ground-based	Space-based
Accuracy	+ Accuracy can be increased by adding beacons - Accuracy severely influenced by ionospheric composition - The achievable accuracy has yet to be studied - The Martian environment prevents the use of groundwaves but allows ionospheric propagation	+ A sub-$10m$ accuracy is achievable - Accuracy severely influenced by ionospheric composition - Long delay to acquire position information with small constellation - The time it takes to obtain a fix depends on latitude.
Coverage	+ Coverage can be tailormade according to mission requirements + No energy, time or resources wasted on providing coverage where there is no need for it	+ Near-global coverage available - It can take a while for enough satellites to be in range to provide position information
Availability	+ Robustness can be increased by providing more beacons + Low beacon complexity will ensure that severe increases in mission cost can be mitigated	+ Robustness can be increased by providing backup satellites - Increasing number of satellites will severely increase mission cost
Ambiguity	- Suffers from ambiguities if not enough beacons are in range	- Suffers from ambiguities if not enough satellites are visible
Mission complexity	+ The cost & complexity of individual beacons is low - Requires ED&L	- Requires precise knowledge of orbits - High unit cost & complexity

time delay required to obtain a fix. The time delay can be decreased by increasing the number of satellites. However, increasing the number of satellites is an expensive endeavour as they have a high unit cost.

Second, in terms of coverage, the availability of global coverage can be considered a disadvantage. Terrestrial users of positioning systems are evenly distributed over the globe, and hence global coverage is a requirement. On Mars, such is not the case. Scientific missions are few and far apart, and global coverage provides little benefit and can mostly be considered a waste of resources. A ground-based positioning system instead allows for a tailormade coverage area which suits the needs of the missions at hand. Nevertheless, while this drawback is true for the current and near-future missions, in the long term Mars may face large scale human colonization. At this point systems which can provide global coverage will become increasingly more desirable.

Conclusion From the discussion it appears that space-based systems have drawbacks that oppose the need for positioning information required by Mars missions. Space-based systems suffer from high time delays when a low number of satellites is used. Furthermore, they provide global coverage instead of coverage that is tailormade to the missions at hand. Hence, a ground-based system is preferred. Due to the advantages of global coverage from

space-based systems, however, they will face increased applicability once the number of missions to Mars grows, for example in the case of colonization of Mars.

For the moment, however, a ground-based system can provide positioning information at a lower mission cost and complexity than a space-based positioning system. However, this statement depends on the assumption that a ground-based system can provide adequate positioning information under all circumstances. This assumption must be validated by testing to what extend positioning information can be provided.

The conclusion of this chapter therefore yields the following hypothesis, which will be tested in the next chapters of this report.

Hypothesis *A ground-based system can provide positioning information required by current and future Mars missions at a competitive precision, availability and coverage compared to a space-based positioning system.*

3 — The Martian Environment

The Martian environment is composed of various elements which influence the propagation of radio signals broadcasted from the surface. These elements can alter various aspects of a signal, such as a reduction in signal strength, a change in signal polarization or a refraction in the direction the signal is travelling. A study on these effects is imperative for the construction of a simulation on signal propagation.

The main environmental influences on signal propagation are the ionosphere, atmosphere and topology. Each of these are studied separately in this chapter, starting with the processes that occur in the ionosphere, descending through the atmosphere and concluding with the topography and surface composition of Mars.

3.1 Ionosphere

The ionosphere around Mars influences radio signals due to its ionized nature. Compared to the ionosphere on Earth, Mars does not have an intrinsic magnetic field shielding it from the solar wind. As a result of the conductive nature of an ionosphere, the interplanetary magnetic field generates a magnetic field around the planet forming a bow shock and diverting the solar wind (Kivelson and Russel, 1995, p.208). Mars receives a weaker solar radiation than Earth, as is further away from the Sun. This is cause for lower electron densities accross the ionosphere. During the nightside there is no solar radiation at all, causing a significant drop in electron number density which results in an electron density profile that is distinctly different from the dayside ionospheric profiles. Typical profiles depicting the difference between the day- and nightside ionosphere are depicted in figure 3.1.

The peak electron density and radio frequency influence the amount of attenuation of radio signals propagating through the various layers of the ionosphere. A decrease in density will cause a decrease in attenuation. Hence, the most severe attenuation will be seen during the dayside ionosphere as the peak electron density is at its highest during the day.

3.1.1 Composition

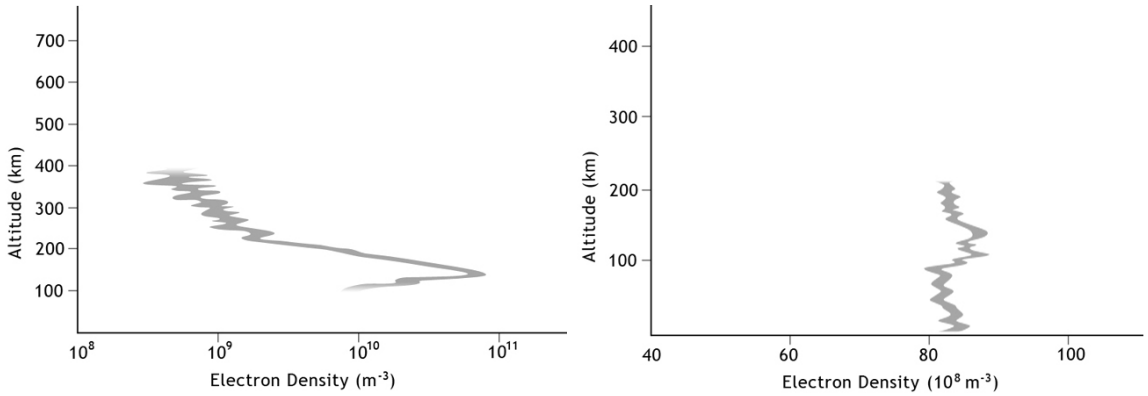
The ionosphere consists of four layers each with different electron density profiles and composition. These layers are defined as follows (Withers, 2011):

- M2 layer, the main ionospheric layer, caused by solar irradiance
- M1 layer, a lesser layer, caused by X-rays
- Meteoric layer, caused by an influx of micrometeorites
- EP layer, a plasma generated by energetic particles

These layers and their impact on the propagation of radio signals are treated in this section.

M2 layer

The most important layer is the M2 layer, which has a peak electron density at $135km$ altitude. The dayside profile can be described by a Chapman profile (Withers, 2009). This



(a) Dayside ionosphere *(b) Nightside ionosphere*

Figure 3.1: Typical electron density profiles for the dayside ionosphere (3.1a) and the nightside ionosphere (3.1b)[1]*. As can be seen from the figures the profile is completely different during the night than during the day. This discrepancy is caused by the absence of solar radiation during the night. In figure 3.1a the peaks for the M1 and M2 layer become apparent as well. The main peak in the profile is due to the M2 layer, while a smaller layer is visible at an altitude slightly above* $100km$*. This smaller peak is from the M1 layer.*

Chapman-like profile is clearly visible in figure 3.1a. Solar extreme ultraviolet is the main cause for photoionization of the CO_2 molecules in this layer. These molecules are the most dominant molecules in the M2 layer. However, they have a short lifetime, as the CO_2 ions interact with oxygen atoms to form O_2^+ ions which have a longer existence than the carbon dioxide ions (Chen et al., 1978; Withers, 2011).

In the M2 layer only the lower frequencies are significantly effected by attenuation(Withers, 2011). In general, attenuation is relatively insignificant and again dependent on the electron peak density. For frequencies below $6MHz$, an attenuation in the order of $1dB$ can be expected.

M1 layer

Slightly below the M2 layer at $110km$ is the M1 layer, which is highly variable in electron density. It is created by the photoionization of solar X-rays and the production of photoelectrons, the latter of which cause more ionizations (Fox, 2004). The photoelectrons are highly energetic and generally cause more ionizations of carbon dioxide molecules. The high variability in electron density can be explained by looking at the origin of solar X-rays. Their intensity depends on the solar activity and thus it varies according to the 11-year solar cycle. Furthermore, solar flares can increase the electron number density in the M1 layer by a factor of two (Mendillo, 2006).

Meteoric layer

Another layer of meteoric origin persists at an altitude of around $80km$ (Molina-Cuberos et al., 2003; Pesnell and Grebowsky, 2000). The presence of this layer was first discovered by the Mars Express orbiter, and it is created by a continuous influx of micrometeoroids coming from the asteroid belt. The ionization is mainly caused by Solar photoionization of

[1] `http://sci.esa.int/mars-express/51056-new-views-of-the-martian-ionosphere/`

metal atoms. This photoionization is the dominant driver for the ionization of the metal ions coming from micrometeoriods because the Martian atmosphere has a low absorptance with respect to the solar ultraviolet radiation. Meteoric material is mainly composed of silicon, magnesium and iron but only the latter two have a significant occurrence in the meteoric layer. This is due to the fact that silicon reacts quickly with dioxide (Molina-Cuberos et al., 2003).

Witasse et al. (2001) and Molina-Cuberos et al. (2003) demonstrated the impact of the meteoric layer on the attenuation of HF radio signals which have frequencies ranging from 1 to $10MHz$. Radio signals propagating through this layer may experience an attenuation of $18dB$ at $9MHz$ and up to $360dB$ at $1.8MHz$. These numbers are in the same order of magnitude as presented by Withers (2011) confirming the severe attenuation. In all cases the electron peak density impacts on the actual severity of the attenuation.

EP layer

The fourth and final layer, also referred to as the *EP* layer, is caused by the influx of energetic particles producing a plasma in the lower ionosphere. The electron density peaks at an altitude of $35km$ and the actual density depends on the solar cycle (Molina-Cuberos, 2002). This layer persists during the nightside ionosphere, because the particles which produce the plasma originate from the cosmos and not just the Sun. According to Withers (2011), the EP layer is the biggest contributor to attenuation for frequencies above $50MHz$. For low frequencies the attenuation is insignificant compared to that in other layers. For example, an attenuation of $1dB$ occurs at only $4MHz$. Here too the attenuation decreases with increasing frequency such that signals with a frequency in the order of $100GHz$ will only experience an attenuation in the order of $10^{-5}dB$ (Withers, 2011).

3.1.2 Frequency limitations

The composition of the various layers limit the working frequencies of radio signals propagating through them. Radio signals propagate through the ionosphere without issue if their frequency is much above the plasma frequency. This frequency determines the maximum frequency for which a signal with a vertical incidence is reflected, and it depends on the electron number density. Signals with frequencies below the plasma frequency do not penetrate the ionosphere, but are reflected instead. This is beneficial for beyond-the-horizon communication. Melnik and Parrot (1999) already suggested working HF frequencies for communication between bases using the ionosphere. Working frequencies between $1MHz$ and $4MHz$ are useful for such communication (Ho et al., 2002; Melnik and Parrot, 1999). The actual frequency used depends on the time of day and the solar activity.

3.2 Atmosphere

The Martian atmosphere interferes with radio signals even though it is weak compared to the Terrestrial atmosphere. Mars' thin atmosphere is primarily composed of carbon dioxide and the average pressure on the surface is only $6mbar$ (Shirley and Fairbridge, 1997), which is generally around 1% of the pressure in the Earth atmosphere at sea level.

Mars only receives half as much sunlight as the Earth due to its large distance from the Sun, resulting in a lower atmospheric temperature. The atmosphere is largely transparent

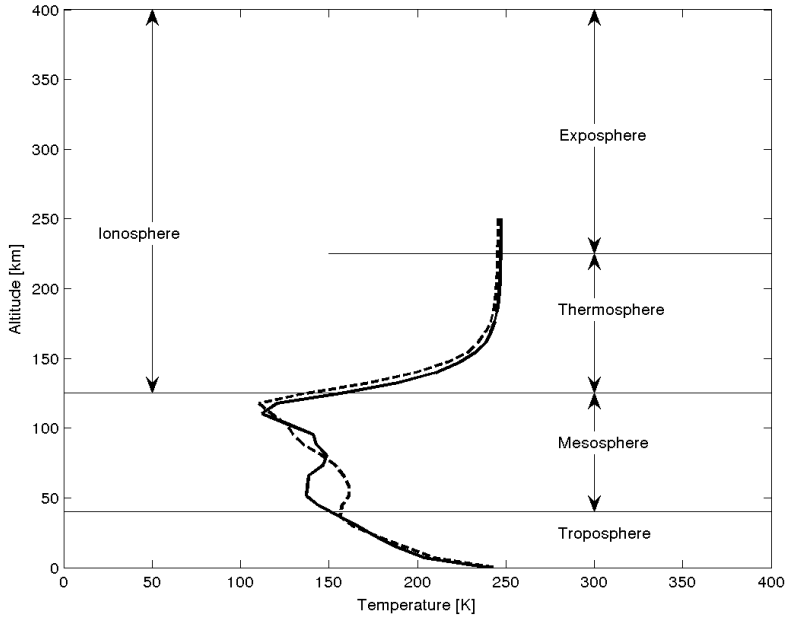


Figure 3.2: Vertical profile of the martian atmosphere (Forget and Millour, 2014)

to solar radiation. Mars has a weak greenhouse effect, which does not help increasing the average temperature. In fact, dust particles have a more significant impact on the temperature as they shield both solar radiation and thermal emissions from the surface (Shirley and Fairbridge, 1997, p.432-436). The temperature profile varies between $150K$ in the lower atmosphere to $215K$ on the surface, as can be seen in figure 3.2. The neutral scale height is a function of temperature and can vary between $10km$ and $15km$. As the altitude of the electron peak density varies with scale height, a change in temperature will ultimately impact the electron density profile as well.

3.2.1 Lower atmosphere

Radio propagation in the lower atmosphere is mainly governed by the effects of refraction and, although radio signals are refracted, this effect occurs to a lesser extend than on Earth. According to Ho et al. (2002), the refractive index of the atmosphere is in fact two orders of magnitude smaller than on Earth, and as a result radio signals will be less affected by atmospheric refraction. This is due to the low atmospheric pressure and the temperature. The radio refractivity depends on the atmospheric pressure, temperature and water vapour pressure (ITU-R P.453-6, 1997). Both an increase in pressure and water vapour will increase the index of refraction and thus rays will experience more severe bending. Temperature on the other hand will decrease the index of refraction.

The effect of water vapour is more significant than previously thought. Early research on the water vapour density concluded that it is almost non-existent, especially compared to the levels on Earth (Shirley and Fairbridge, 1997). However, recent research by Maltagliati et al. (2011) has shown that the water vapour levels are significantly higher than previously thought. In the upper troposphere, the water vapour levels can reach mixing rations of up to $20ppmv$ (*parts per million by volume*), where the highest ratios are found in the north-

ern polar region.

The index of refraction changes in the troposphere only and stays constant in the mesosphere. This profile is caused by the dependence of the index of refraction on the temperature and pressure (ITU-R P.453-6, 1997). By very definition, the troposphere is that part of the lower atmosphere where the temperature decreases. Figure 3.2 shows that this decrease occurs up to an altitude of about $40km$, at which point the temperature is around $150K$. Therefore, the index of refraction decreases with an increase in altitude, but stay constant in the mesosphere.

On Earth, signals can get trapped if their curvature follows the curvature of the planet. This effect is called trapping, but will almost not occur on Mars because the refractive index is too low to cause any significant curvature. Ho et al. (2001) calculated that the angle between a radio signal and the horizontal needs to be less than $0.3\,\mathrm{deg}$ for trapping to occur.

3.2.2 Seasons

The high eccentricity of Mars' orbit affects the influx of solar radiation as its intensity gradually decreases with increasing dinstance. This in turn impacts the difference between seasons in the northern hemisphere and their southern counterparts (Morgan et al., 2008; Nielsen et al., 2007). The southern hemisphere generally experiences long winters while the northern part experiences long summers. As a result the electron number density in the ionospheric layer will be lower during winters in the southern hemisphere. This decrease influences both the attenuation and usable frequency, as described in section 3.1.

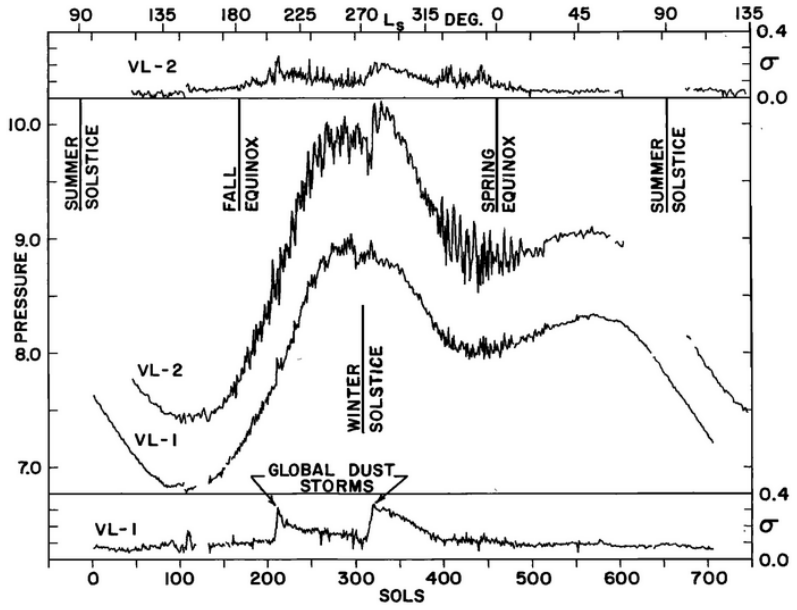


Figure 3.3: The annual variation in pressure in the northern hemisphere as measured by the two viking landers (Hess et al., 1980).

The seasonal variations globally impact the pressure and hence the nature of radio signal propagations. The seasonal pressure reaches a minimum during the summer in the northern hemisphere. These variations were first measured by the two viking landers (Hess et al., 1980). A depiction of the variation can be seen in figure 3.3. Later observations done by

the Mars Pathfinder confirmed these findings (Schofield, 1997). The Pathfinder mission measured minima in pressure which occured 6 days later than the minima observed by the Viking missions. The annual fluctuations in pressure affect the refractivity experienced by radio signals, as described in section 3.2.1.

3.2.3 Dust storms

Dust storms occur frequently on the red planet but they only cause a minor attenuation to radio signals. Dust storms have first been reported by the Viking missions (Shirley and Fairbridge, 1997). Smaller dust storms occur more often, but a global dust storm occurred twice during the Viking missions.

The temperatures in the lower atmosphere increase during a dust storm and hence the density increases. Essentially, dust storms cause an expansion of the entire lower atmosphere, which also impacts the altitude of the peak electron density as the profiles are shifted upwards as well (Bougher et al., 2001). The altitude of the peak electron density increases by as much as $30km$ (Wang and Nielsen, 2003), which will severely impact the behavior of the ionosphere on radio signals. Therefore, the upward shift of the ionosphere due to dust storms must be taken into account when analyzing the propagation of these signals.

The attenuation due to a dust storm depends on the frequency and solar zenith angle of a radio signal as well as the material properties of the particles in the dust storm (Smith and Flock, 1986). Both an increase in frequency and solar zenith angle will increase the attenuation. Still, Smith and Flock (1986) showed that the attenuation is only $1.3dB$ even in the worst case for a $32GHz$ radio signal tangential to the Martian surface. This implies that the attenuation will be even less for lower frequencies and lower solar zenith angles. This is in contrast with the suggestion made by Annis (1987) based on optical depth measurements. Annis (1987) doubts the successful propagation of radio signals without going into detail.

The optical depth of dust, which is constantly present on the red planet, is a dimensionless measure for attenuation of a radio signal. Aerosols and dust have an optical depth of up to 0.5 (Ho et al., 2002) and it varies with the season (Smith, 2009). The optical depth is at it highest when Mars is closer to the Sun. Lower optical depths, in the order of $0 - 0.1$ have been detected during the northern winter, at which time Mars is at its apohelion. The optical depth increases severely during dust storms. The Viking landers measured maxima of 9 during and shortly after dust storms in which they were encompassed (Annis, 1987; Pollack et al., 1979).

3.2.4 Clouds and fog

Clouds are a common feature on Mars, even though the atmosphere is very dry compared to Earth. They are composed of water ice and until recently it has been assumed that the clouds contain little amount of water. A recent study by Maltagliati et al. (2011) however has demonstrated that supersaturation of clouds in fact occurs more frequently and more severely than on Earth. In a supersaturated state the water vapor is not in equilibrium and will either precipitate or be lifted to higher altitudes. It is generally assumed that precipitation on Mars by water does not occur. The water vapor rising through the atmosphere will be affected by photo dissociation, turning the water into its constituents H and O. These can subsequently escape the planet, and as such influencing the rate at which Mars loses water.

The occurrence of clouds has a minor effect on radio signals, as the scattering caused by

them causes some attenuation. Clouds are more common in the north (Annis, 1987; Shirley and Fairbridge, 1997). They are seasonally dependent, the most dense clouds occurring during northern spring and summer. The severity of the attenuation correlates to the optical depth. The optical depth is a measure of transmittance through a gas, which indicates how much of a signal passing through that gas is absorbed. For clouds, the optical depth is around 1.0 (Annis, 1987). Dense clouds have a slightly higher optical depth, thus causing more attenuation. Similarly, the water ice optical depth depends on the season and can obtain values up to 0.2 (Smith, 2009). Still, at frequencies of around $32GHz$, the attenuation caused is expected to be in the order of $0.01dB$ (Ho et al., 2002). The attenuation decreases as the frequency decreases. Therefore, the effects of attenuation can be safely ignored for the frequencies of interest.

3.3 Topography

Radio signals are affected by the presence of mountains and valleys, as well as the very regolith the planet is made of. Not only that, but local magnetic fields influence radio signals as they propagate through the atmosphere and ionosphere.

3.3.1 Topology

The topology of the red planet can alter radio signals due to the shape of the terrain and the composition of the terrain, causing multipath effects and diffraction. Multipath effects are caused when radio signals are reflected or diffracted by surface features, thusly reaching a receiver at a time delay with respect to an unobstructed signal. These effects can occur practically anywhere on the planet. They are greatly amplified in regions with a strong diversity in topology. The main example of such a region is the Valles Marineris region, which consists of vast canyons with a depth up to $9km$ (Shirley and Fairbridge, 1997). Rovers located inside a canyon or valley will face an obvious decrease in direct line-of-sight communication opportunities with beacons or satellites. Communication is nonetheless possible as signals are diffracted at canyon edges, thereby covering areas in the canyons that are not covered by direct line-of-sigth signals (Ho et al., 2002).

Data from the Mars Global Surveyor has shown that canyons and valleys are subject to an increase in optical depth by as much as 0.5 depending on canyon depth (Ivanov and Muhleman, 1998). This affects the reception of radio signals negatively and can even lead to a complete loss of signal due to the local opaque atmosphere (Ivanov and Muhleman, 1998).

The roughness of the regolith in canyons will diminish the effects of multipath propagation (Ho et al., 2002). As such multipath effects will only play a significant role in regions with a relatively low surface roughness. The Martian poles, mostly covered in carbon dioxide (Shirley and Fairbridge, 1997), are an example of such a region. Some canyon features are apparent in the poles, and here multipath effects will play a role too(Ho et al., 2002). The advantage of multipath effects is that a receiver will be able to receive radio signals that would otherwise be obstructed by the terrain. However, signals are greatly attenuated.

3.3.2 Crustal magnetic field

The main impact of a magnetic field on a radio signal is the faraday rotation caused by the presence of a magnetic field (Safaeinili et al., 2003). Without its presence, the refractive index in the ionosphere is changed isotropically. When a magnetic field is present however, radio signals with different polarizations will be refracted differently, causing group delays.

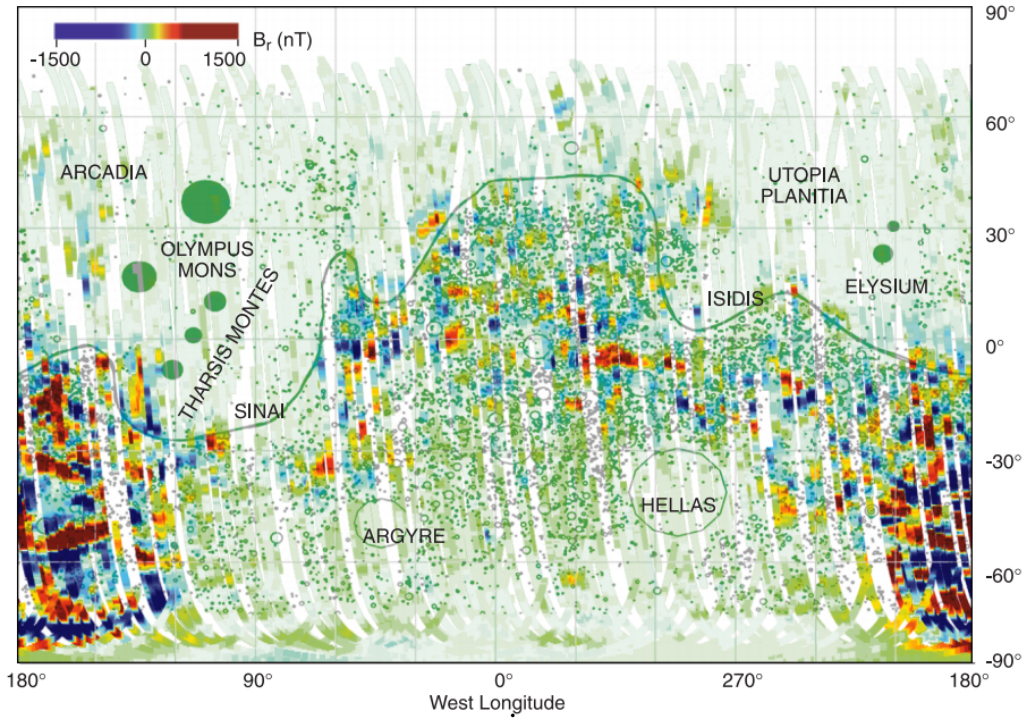


Figure 3.4: The vertical intensity of local crustal fields. The map shows high intensities in the southern hemisphere while the northern hemisphere is devoid of any significant magnetic field (Acuña, 1999)

The Martian magnetic field varies from the magnetic field on Earth in the sense that no intrinsic magnetic field is present on Mars. Instead, local crustal fields are distributed mainly over the southern hemisphere. Initial models suggested that Mars had the necessary conditions to support an intrinsic magnetic field. The rotation rate of the planet is adequate for the generation of a dipole moment, and the accretion of the planet supplied sufficient heat as well (Shirley and Fairbridge, 1997). Recent findings suggest that a series of giant asteroid impacts in Mars' early life were the cause of shutting down the electric dynamo that generates the magnetic field (Roberts et al., 2009). What is left of the magnetic field are local crustal fields which consequently only locally influence the propagation of radio signals.

Data from the Mars Global Surveyor has shown that especially the southern hemisphere is covered in local crustal fields, as depicted in figure 3.4. The fields can sometimes have a magnetic field intensity two orders of magnitude higher than the average intensity (Acuña et al., 2001). This causes the local fields to function as a significant protective layer shielding the atmosphere from eroding due to the solar wind. Therefore, these local fields are a main contributor to radio signal refraction and group delays (Safaeinili et al., 2003).

4 — Environmental Model

A mathematical description of the Martian environment is necessary in order to simulate the behavior of radio signals, propagating through it. Radio signals transport energy from one point to another while being subject to losses caused by the environment in which they are propagating. This study focuses on the propagation through both a neutral and a plasma environment. The interaction of radio signals with a neutral environment causes absorption of energy and thus loss of signal strength. Moreover, the change in particle density causes light to refract. Radio signals can be treated as light and will refract in a neutral environment. Comparable effects occur when the signal is suspended in a plasma. The interaction with the plasma causes the signal to refract and absorb. In order to determine how much absorption and refraction is caused, the composition of the plasma needs to be described.

This environment can be approximated by a mathematical model which consists of the most important features that Mars has to offer. As was found in section 3.2, the refraction caused by the atmosphere is relatively insignificant. However, it will be included in the model for completeness. Atmospheric absorption can be severe, especially if a dust storm is present. The plasma environment is the most prominent contributor to refraction and absorption of radio signals. Starting from Maxwells' equations, their behavior can be studied so that a quantitative approximation can be obtained in relation to the direction the signal is propagating to and how much energy it is carrying.

This chapter provides a mathematical approximation for all features that somehow influence the propagation of radio signals. A qualitative discussion has already been given in chapter 3. Here, the discussion is continued by describing the equations that construct the environment. First, the basic properties of an ionosphere are given in section 4.1. Subsequently, radio signal geometry and absorption models are laid out in the remainder of this chapter. Section 4.2 deals with the geometric properties of a signal as it propagates through the ionosphere. The geometric representation of a radio signal is referred to as a *ray* in the remainder of this work. A ray is described as a vector with a certain frequency. The direction of a ray itself is determined by the way in which it refracts and reflects, as described in section 4.2.1. The iris effect (section 4.2.2) is a convenient metric to express the outer bounds in which rays can propagate. Ultimately, section 4.3 expresses the absorption faced by the signals. As will be seen, absorption levels are experienced differently in the ionosphere and the atmosphere.

4.1 Composition

The ionosphere can be described and approximated by a couple of parameters. First, the frequency at which the plasma oscillates, is a main influential factor for the determination of both refraction and absorption. Hence, this frequency is described in section 4.1.1. Subsequently, the modeling of the electron density is described in section 4.1.2. This naturally leads to the question of finding the total electron content in a vertical profile, which is characterized in section 4.1.3. In an ionized medium like the ionosphere, the collision frequency between neutrals and electrons causes absorption. An analytical expression is derived in

section 4.1.4 so that the magnitude of absorption can be calculated.

4.1.1 Plasma frequency

The plasma frequency is the natural frequency in an ionospheric plasma, independent of the presence of a magnetic field. In the discussion in this research it is assumed the plasma is a cold plasma, and as such the thermal motion is ignored. This leads to the definition of the plasma frequency, ω_p, which depends purely on the electron number density N_e, the electron charge e and mass m_e and the permittivity of free space ϵ (Kivelson and Russel, 1995):

$$\omega_p^2 = \frac{N_e e^2}{m_e \epsilon} \tag{4.1}$$

The unit of the plasma frequency in equation 4.1 is $rads^{-1}$.

The plasma frequency is an important factor in determining radio signal dispersion and refraction. As will demonstrated in a later section, the behavior of a radio signal in an unmagnetized plasma depends on its relation to the plasma frequency. When the angular frequency is above the plasma frequency, so $\omega > \omega_p$, the signal is refracted. In the case of $\omega < \omega_p$, an imaginary refraction persists causing the radio signal to decay and cease propagation. An intermediate case exists where $\omega = \omega_p$. If this situation is encountered during radio signal propagation, the radio signal will reflect instead. As the plasma frequency in a constant environment is essentially a function of the electron number density, it is this value which ultimately dictates the behavior of a radio signal. Therefore, a formulation of the electron number density is required.

4.1.2 Electron number density

The electron number density in an ionosphere can be modeled using Chapman theory (Kivelson and Russel, 1995). A Chapman model relates the ion production in the ionosphere to the altitude. This process causes molecules to be disbanded into ions and electrons. The model is a reasonable simplification under the following assumptions (Kelso, 1964; Kivelson and Russel, 1995):

- The atmosphere consists of a single gas
- Neutral density decreases exponentially with increasing altitude
- Atmospheric absorption is proportional to the density
- The neutral scale height H_n is constant

From the assumptions it logically follows that there is some maximum production rate as the production rate depends on both the neutral density and the solar radiation. The electron number density can be derived from the analysis of the rate of ion production Q with respect to the maximum production rate Q_m. The Chapman production function states (Kivelson and Russel, 1995):

$$Q = Q_m \exp\left[1 + \frac{h_m - h}{H_n} - \exp\left(\frac{h_m - h}{H_n}\right)\right] \tag{4.2}$$

Where H_n is the neutral scale height and h_m the altitude at which the maximum production rate occurs. Equation 4.2 can be simplified by letting:

$$y = \frac{h - h_m}{H_n} \tag{4.3}$$

So far, the effects of the SZA have been neglected. In fact, the production rate is related to the SZA, denoted χ, of the production rate Q_{mo} at the subsolar point. Recall that the the SZA is the central angle to the subsolar point. Introducing this angular dependency gives:

$$Q_m = Q_{mo}\cos(\chi) \tag{4.4}$$
$$h_m = h_{mo} + \log\left(\sec(\chi)\right) \tag{4.5}$$
$$Q = Q_{mo}\exp\left[1 - y - \sec(\chi)\exp(-y)\right] \tag{4.6}$$

The discussion so far focused on the total production rate. The normal continuity equation for the electron density equals $\frac{\partial N_e}{\partial t} = Q - L$, where the loss rate L is assumed to be a product of the recombination coefficient α and the electron density (Kivelson and Russel, 1995, p.193):

$$L = \alpha N_e^2 \tag{4.7}$$

In a steady state environment, the electron density adheres to the continuity equilibrium:

$$\frac{\partial N_e}{\partial t} = Q - L = 0 \tag{4.8}$$

Therefore, the electron density can be related to the production rate:

$$N_e = \sqrt{Q/\alpha} \tag{4.9}$$

Equation 4.2 can be rewritten for the electron density wit the use of equation 4.9:

$$N_e = N_{e,mo}\exp\left[\frac{1}{2}\left(1 - y - \sec(\chi)\exp(-y)\right)\right] \tag{4.10}$$

Equation 4.5 and 4.10 reveal that the peak electron density, and the altitude at which it occurs, is ultimately a function of SZA .

It should be pointed out that the validity of a Chapman-like electron density profile only holds during daytime, that is SZA $< 70°$. During the nighttime, the electron density is significantly lower and does not follow a Chapman profile (Němec et al., 2010). Instead, a profile with constant electron density is better suited for an approximation of the electron density in an ionospheric layer (Withers et al., 2012).

The Chapman function from equation 4.10 ignores temporal fluctuations in the ionosphere, reducing the electron number density to merely a function of altitude and SZA . This is a valid simplification for this work, considering that the fluctuations in the ionosphere are on a timescale several orders of magnitude larger than the propagation time of radio signals. Hence, fluctuations are ignored. Radio signals, traversing at the speed of light, take $1ms$ to propagate $300km$. The propagation distances relevant in this study are in the order of hundreds of kilometers. However, the timeframe in which electron densities in the ionosphere fluctuate is in the order of seconds to minutes (Duru et al., 2008). As this timeframe is several orders of magnitudes higher than the timescale relevant for this study, it is assumed that the ionosphere stays constant during the propagation duration of a radio signal.

4.1.3 Total Electron Content

The total electron content, or *TEC*, describes the total number of electrons which are within a vertical column straight through the ionospheric layer. It is an important metric as it influences the severity of phase and frequency variations experienced by a radio signal. The TEC

can be found by integrating equation 4.10:

$$TEC = \int N_e dh \tag{4.11}$$

The variations in solar intensity and SZA will change the TEC , as the TEC is basically proportional to the electron number density and the layer height. Hence, the phase and frequency variations a radio signal encounters also vary with solar intensity and SZA, as will be seen in a later section.

4.1.4 Collision frequency

Collisions between electrons and neutral particles cause a decay on radio signals as they propagate through an ionized medium (Davies, 1969). The quantitative measure which describes this effect is the electron-neutral collision frequency ν. The intensity of the collision frequency is mainly defined by the composition of neutrals in an atmosphere. In the case of Mars, the atmosphere consists largely of CO_2[1]. Assuming an atmosphere dominated by CO_2, Schunk and Nagy (2009) define the collision frequency ν_{CO2} as a dependency of the number density n_{CO2}, in el/cm^{-3}, and the electron temperature T_e, in Kelvin:

$$\nu_{CO2} = 3.68 \cdot 10^{-8} n_{CO2} \left(1 + 4.1 \cdot 10^{-11} |4500 - T_e|^{2.93}\right) \; [Hz] \tag{4.12}$$

Equation 4.12 can only be solved if the electron temperature and the neutral number density are known. The neutral number density, indicated by $n_{CO2,s}$, is the highest at the surface. It decreases exponentially with altitude h assuming a scale height H_n Kelso (1964):

$$n_{CO2} = n_{CO2,s} \cdot e^{\frac{-h}{H_n}} \tag{4.13}$$

The electron temperature varies with Solar irradiance and increases with increasing altitude (Withers et al., 2014). Simultaneously, research from Withers et al. (2014) has shown that the electron temperature matches the neutral temperature at lower altitudes. Assuming a constant neutral temperature of $180K$, equation 4.12 reduces to:

$$\nu_{CO2} = 1.0436 \cdot 10^{-7} n_{CO2} \tag{4.14}$$

$$= 1.0436 \cdot 10^{-7} n_{CO2,s} \cdot e^{\frac{-h}{H_n}} \tag{4.15}$$

The assumption of a constant neutral temperature will ultimately impact the magnitude of absorption which is encountered by signals. Therefore, this assumption will be validated by comparing the simulated absorption with measured absorption levels in chapter 6.

4.2 Geometry

The geometry of a radio signal defines what path it follows through a certain environment. The equations required to determine this path are derived in section 4.2.1. Radio signals do not necessarily always return to the planetary surface. One major effect that causes this is shown in section 4.2.2.

[1] http://nssdc.gsfc.nasa.gov/planetary/factsheet/marsfact.html

4.2.1 Refractive index

The dispersion of a radio signal in a plasma follows from the *Appleton-Hartree Dispersion Equation*. This equation applies to radio signals propagating in a cold magnetized plasma. It is a recommended standard formulation (Kelso, 1964; Kivelson and Russel, 1995) and will therefore be used as the basis to derive equations for both the index of refraction and absorption. The full dispersion equation, defined in (Kivelson and Russel, 1995), has the following form:

$$n_c^2 = 1 - \frac{X}{(1 - iZ) - \frac{Y_\perp^2}{2(1-X-iZ)} \pm \sqrt{\frac{Y_\perp^4}{4(1-X-iZ)^2}}} \tag{4.16}$$

A couple of substitute variables X, Y and Z are used in equation 4.16 to enhance readability. These are defined as follows:

$$X = \frac{\omega_p^2}{\omega^2} \tag{4.17}$$

$$Y = \frac{\Omega_c}{\omega} \tag{4.18}$$

$$Y_\parallel = \frac{\omega_\parallel}{\omega} \tag{4.19}$$

$$Y_\perp = \frac{\omega_\perp}{\omega} \tag{4.20}$$

$$Z = \frac{\nu}{\omega} \tag{4.21}$$

In the above equations Ω_c represents the angular gyrofrequency of a particle suspended in a magnetic field. Similarly, the angular frequencies for the parallel and perpendicular components of the magnetic field, measured with respect to the direction of propagation, are represented by equations 4.19 and 4.20 respectively.

The composition of Mars allows for a few simplifications in the Appleton-Hartree relationship. Mars has no intrinsic magnetic field and the local fields which do occur are two to three orders of magnitude lower in field strength than the magnetic dipole of the Earth (Acuña et al., 2001). Therefore, the magnetic field effects on Mars are ignored and an isotropic ionosphere assumed. As a result, it holds that $\{Y, Y_\parallel, Y_\perp\} << \{X, Z\}$. This reduces the complexity of equation 4.16 enormously:

$$n_c^2 = 1 - \frac{X}{1 - iZ} \tag{4.22}$$

Continuing from the simplified Appleton-Hartree relationship, the complex index of refraction can be rewritten in terms of the real index of refraction and the index of absorption:

$$n_c^2 = (\mu - i\xi)^2 \tag{4.23}$$

The real part, μ, in equation 4.23 represents the index of refraction and the imaginary part ξ the index of absorption. Equations for the real and imaginary parts can be found by combining equations 4.22 and 4.23. Then, the result can be solved for the real and imaginary part

separately:

$$\mu^2 - \xi^2 = \mathrm{Re}((\mu - i\xi)^2) \quad (4.24)$$
$$= 1 - X \quad (4.25)$$
$$2\mu\xi = \mathrm{Im}((\mu - i\xi)^2) \quad (4.26)$$
$$= \frac{XZ}{1+Z^2} \quad (4.27)$$

The approximate form of the refractive index in a non-magnetic environment follows from the combination of equations 4.25 and 4.17(Davies, 1969):

$$\mu^2 = 1 - \frac{\omega_p^2}{\omega^2} \quad (4.28)$$

Equation 4.28 demonstrates the conspicuous limitations of refraction in an unmagnetized plasma. The refractive index becomes imaginary as soon as $\omega_p > \omega$. At this point propagation through the ionosphere is not possible and the radio signal ceases to exist. However, a special case exists where $\omega_p = \omega$. Radio signals are *reflected* when their frequencies equal the plasma frequency, as was shown in section 4.1.1.

4.2.2 Iris effect

The iris effect dictates whether a ray penetrates the ionosphere, or is refracted by it instead. Whichever of the two occurs, depends on the angle of the ray with respect to the zenith. Vertical rays, thusly launched in the direction of the zenith, penetrate the ionosphere if their frequency is above the plasma frequency. An oblique ray with a small zenith angle might still penetrate. A minimum zenith angle exists, for which rays no longer penetrate the ionosphere, but are refracted instead. It can therefore be used to validate the propagation of simulated rays. The minimum zenith angle is referred to as the iris angle ϕ_c, and it is measured with respect to the surface normal. The iris angle depends on both the signal and

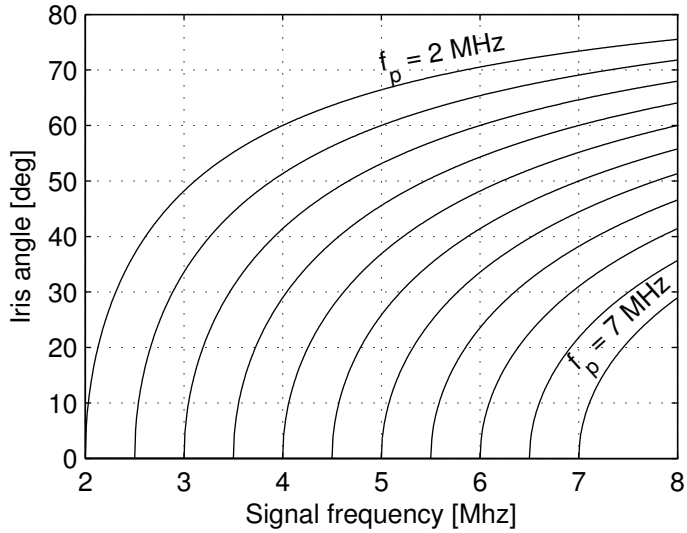


Figure 4.1: The iris angle as a function of the signal frequency.

plasma frequencies (Kelso, 1964):

$$\phi_c = sin^{-1}\left(\sqrt{1-\left(\frac{f_p}{f}\right)^2}\right) \tag{4.29}$$

Signals with an iris angle higher than the critical frequency are refracted by the ionosphere, whereas signals with a lower angle disappear into outer space. Therefore, this effect is called the *iris effect*: a hole, or iris, in the ionosphere exist above some signal origin, through which rays pass into outer space.

The critical angle, defined in equation 4.29, describes the boundary between two types of rays, those who are refracted and those who penetrate the ionosphere. Refracted rays are those whose parallel component of the signal frequency is below the plasma frequency.

The impact of equation 4.29 has been plotted in figure 4.1. Obviously, the iris angle ϕ_c depends on the relationship between the plasma frequency and the ray frequency. The plasma frequency rises with the rise of the electron number density in the ionosphere. According to equation 4.29, this causes the iris angle to decrease. Thus, an increase of plasma frequency yields a smaller iris angle for the same signal frequency. Conversely, an increase in signal frequency causes an increase in iris angle. Signals with a higher frequency are subject to a larger iris in the ionosphere, for which they are not refracted. Note that the iris effect only holds for signal frequencies above the plasma frequency, see also equation 4.29. Signals with a frequency below the plasma frequency are always reflected.

4.3 Absorption

Radio signals are attenuated as they propagate through the atmosphere and ionosphere. The attenuation consists of both absorption and scattering. Scattering is divided into incoherent scattering and coherent scattering. It is assumed that both types of scattering are insignificant for this work, as that is the usual practice (Nielsen et al., 2007). Hence, in the rest of this work, the term absorption is used. The absorption is caused by the collision between the signal and environmental particles. Upon collision, some of the energy of the signal is transferred to neutral particles in the environment, and eventually lost as thermal energy. The energy in the signal decreases until it eventually reaches zero, at which point the signal ceases to propagate. Nevertheless, it is imperative to study the levels of absorption, faced by signals, in order to determine how far they can propagate. Apart from the issue of propagation, the magnitude of absorption in itself contains valuable information. The energy which is ultimately received by an observer can provide a measure of the length of the path a radio signal has traveled. This information can be used to derive the distance between the observer and a ground beacon.

This work uses the unit decibel to express magnitudes of absorption, and hence, a quick overview of this unit is given. The absorption in decibels is measured as the ratio between the transmitted power P_t and the received power P_r. This ratio is measured on a logarithmic scale, and it is expressed as follows:

$$N = 10\log_{10}\left(\frac{P_r}{P_t}\right) \tag{4.30}$$

In a line-of-sight (LOS) environment, the received power with respect to the transmitted power decreases with distance. As will be seen in a later chapter in this work, the received

power does not follow the same behavior in the ionosphere, as it is a typical non-line-of-sight (NLOS) environment.

4.3.1 Ionosphere

The discussion on the refractive index in section 4.2.1 can be used to derive an expression for the absorption of a radio signal in an unmagnetized, cold plasma. The Appleton-Hartree relation yields a complex index of refraction which contains information about both the refraction and the absorption of a radio signal. The relation between these values is denoted in equation 4.23. The index of absorption, ξ, follows from the complex part of the complex index of refraction. Solving equation 4.27 for ξ yields the following expression for the index of absorption:

$$\xi = \frac{1}{2\mu}\frac{XZ}{1+Z^2} \tag{4.31}$$

The absorption per unit length is a function of the index of absorption, the radio signal angular frequency and the speed of light (Davies, 1969):

$$\kappa = \frac{\omega}{c}\xi \tag{4.32}$$

Combining equations 4.31 and 4.32 together with equations 4.17 to 4.21 yields a complete numerical expression for the absorption per unit length:

$$\kappa = \frac{\omega}{2c\mu}\frac{XZ}{1+Z^2} \tag{4.33}$$

$$= \frac{e^2}{2c\mu m_e \epsilon}\frac{N_e\nu}{\omega^2+\nu^2} \tag{4.34}$$

$$\approx 4.61\cdot 10^{-5}\frac{N_e\nu}{\omega^2+\nu^2}\,[dB/m] \tag{4.35}$$

In equation 4.35, the electron number density N_e is expressed in $el\cdot m^{-3}$, and the collision frequency and plasma frequency in Hz and $rad\cdot Hz$, respectively.

There are a number of limit cases which can be distinguished from equation 4.34. These cases give some insight in the behavior of the absorption for various radio frequencies (Davies, 1969; Withers, 2011). First, there is the low-frequency limit, which occurs when a radio signal propagates through a region where the collision frequency is the dominating factor, such that $\omega^2 << \nu^2$. In this case, equation 4.35 can be simplified to:

$$\kappa \approx 4.61\cdot 10^{-5}\frac{N_e}{\nu} \tag{4.36}$$

The absorption increases with decreasing collision frequency. Considering that the collision frequency decreases with increasing altitude (see section 4.1.4), it can be deduced that the absorption for the low-frequency limit will increase with an increase in altitude. This behavior can be found in the lower region of the ionosphere, below the layer peak (Davies, 1969). Second, there is the high-frequency limit, which occurs in regions where the radio frequency is significantly higher than the collision frequency, or $\omega^2 >> \nu^2$. Hence, equation 4.35 reduces to:

$$\kappa \approx 4.61\cdot 10^{-5}\frac{N_e\nu}{\omega^2} \tag{4.37}$$

Equation 4.37 demonstrates that the absorption is inversely proportional to the radio frequency in this case. The high-frequency limit occurs in the upper regions of the ionosphere, above the layer peak, because the collision frequency decreases with increasing altitudes. Last, there is the intermediate case, for which the radio frequency and collision frequency are comparable. The edge cases show that the absorption increases in the lower regions, and decreases in the upper regions of the ionosphere. Hence, a maximum must exist in between these edge cases. A derivation of equation 4.35 shows that the maximum absorption per unit length is obtained when the angular radio frequency equals the collision frequency. This is a situation which is preferably avoided.

The absorption per unit length can be numerically integrated to find the full absorption of a radio signal. The absorption per unit length, as obtained from equation 4.35, is expressed in dB/m. The total absorption caused by one ionospheric layer is found from integration of the absorption over the path length s:

$$A = \int_0^s \kappa ds \tag{4.38}$$

The total absorption in equation 4.38 naturally has the unit dB. Now, consider the path length s to be equal to the path which a radio signal takes through a certain ionospheric layer with fixed with and fixed ionospheric properties. The total absorption then simply amounts to the superposition of the absorption encountered in a certain composition of ionospheric layers. This approach allows for a fast numerical integration of the absorption, ultimately yielding the total absorption which a radio signal experiences over its entire path through the ionosphere. Such an approach will be used in subsequent chapters as a numerical model for the absorption in an ionospheric environment.

4.3.2 Atmosphere

The atmospheric absorption is mainly governed by the absorption caused by dust storms. Chapter 3 already demonstrated that other effects cause absorptions insignificant enough to be ignored, and therefore the discussion will focus on dust storms.

The absorption depends on the frequency of collision between the radio signal and dust particles. Larger and more numerous particles cause more absorption. Furthermore, the wavelength of the radio signal influences to what extend it interacts with the dust particles. Generally, a longer wavelength causes a decrease in absorption. The absorption per unit length can thus be expressed as follows (Ho et al., 2002):

$$\kappa_d = c_d \frac{\bar{r_d}\tau}{\lambda} \left[\frac{3\epsilon''}{(\epsilon' + 2)^2 + \epsilon''^2} \right] \tag{4.39}$$

$$= c_d \frac{c\bar{r_d}\tau}{f} \left[\frac{3\epsilon''}{(\epsilon' + 2)^2 + \epsilon''^2} \right] \tag{4.40}$$

In equation 4.39, the size of the dust particles is denoted by $\bar{r_d}$, expressed as the mean particle radius. The optical depth is a measure of transmittance through a layer of dust and in here it is represented by τ. A higher optical depth means a more opaque dust layer and thus more absorption by the particles. The expressions ϵ' and ϵ'' follow from the general expression of permittivity where $\epsilon = \epsilon' + i\epsilon''$. Thus, ϵ' and ϵ'' represent the real and imaginary parts of the permittivity respectively. The radio signal is expressed by its wavelength, λ, and immediately replaced by its frequency in equation 4.40, where $\lambda = f/c$. Lastly c_d represents a constant, with $c_d = 54.62 \cdot 10^{-3}$ (Ho et al., 2002).

A distinct feature of equation 4.40 is its dependence on the material properties of the dust. Hence, this equation can only be evaluated if the composition of the dust particles is known.

Table 4.1: Properties of Martian dust particles for HF frequencies

Property	Value	Source
Mean particle radius	$\bar{r_d} = 10\mu m \pm 2.0$	Ho et al. (2002)
Optical depth	$\tau = 3.7$ to 9	Pollack et al. (1979)
Permittivity	$\epsilon' = 5.0 \pm 0.1$	Heggy et al. (2007)
Permittivity	$\epsilon'' = 0.06 \pm 5 \cdot 10^{-3}$	Heggy et al. (2007)

Table 4.1 represents the properties of Martian dust particles for the frequencies of interest. The permittivity varies with the frequency (Heggy et al., 2007) and is therefore displayed for HF frequencies. Using these properties, the expression of the absorption from equation 4.40 can be greatly simplified:

$$\kappa_d = \frac{451.42}{f} \tag{4.41}$$

It is obvious from equation 4.41 that, for a given dust storm environment, the absorption per unit length purely depends on the radio signal frequency.

The total absorption through the atmosphere can be found by integrating for κ_d akin the procedure leading to equation 4.38. By further taking into account the angle of the ray incident to an atmospheric layer, denoted by θ_i, the following equation is obtained:

$$A = \int_0^s \kappa ds \tag{4.42}$$

$$= 451.42 \int_0^{h_d} cos(\theta_i) \frac{ds}{f} \tag{4.43}$$

As the absorption from dust storms obviously only applies to the altitudes where the dust storms are active, the integral in equation 4.43 has been set to cover the part of the path traveled in the dust storm. Assuming the top altitude of dust storm activities is at h_d, the path a radio signal travels through the dust storm is equal to $s = h_d/cos(\theta_i)$. This notation has been reflected in equation 4.43.

5 — Method of Ionospheric Simulation

The influence of the Martian environment upon the propagation of radio signals can be determined by a numerical simulation. Doing so allows gathering propagation data which can be used to predict the precision that a ground-based beacon system can provide. In this chapter, a method is proposed to gather the data needed to obtain a quantitative prediction on the precision and subsequently test the hypothesis. The method involves the development of a simulation[1] which determines the propagation of individual radio signals, or *rays*, which interact with the environment. The interaction with the environment causes the rays to refract, reflect, attenuate and distort in other ways. Ultimately, these rays are picked up by some mobile station and their properties are used to estimate a position. How well a signal is received by a mobile station dictates how well this mobile station can estimate its position. In other words, the precision of the position estimate results from the physical properties of the received signal. This information in turn is valuable to the testing of the hypothesis.

This chapter provides a numerical simulation tool to study radio signal propagation by means of a Monte-Carlo ray-tracing simulation. The general method derives from the technique of ray-tracing as applied in the field of physics. Section 5.1 expresses the ray-tracing approach in more detail. Mathematical models for the refraction and reflection of a ray are presented in section 5.2. These methods are used in the simulation to numerically approach the effects of the Martian environment upon a ray. The chapter continues with describing how a Monte-Carlo approach is used to model dispersion in the results, caused by ionospheric deviations, in section 5.3. Finally, an architectural overview of the software simulation is given in section 5.4, which combines the technique of ray-tracing with the mathematical models of the Martian environment. Moreover, section 5.4.6 treats how the calculations in the simulations are verified. A user and installation manual of the simulation are given in appendix C.

5.1 Ray-tracing

Ray-tracing allows the projection of the path of a particle through a system consisting of elements with various properties. Radio signals can be described as a ray and as such ray-tracing can be used to trace the propagation of a radio signal to an environment.

In this section, the theory of ray-tracing a radio signal is described. Subsequently, an algorithm with a ray-tracing implementation is shown which will serve as the base of the numerical simulation. The section ends with a description on the methods of approximation used by this ray-tracing method, such that a stepwise simulation is possible.

5.1.1 Technique

It is important to understand the concepts of ray-tracing first before this technique can be applied to tracing the propagation of a radio signal. In the subsequent text, radio signals are referred to as *rays* which propagate through a certain *scene*. The scene is a numerical interpretation of the Martian environment in this case.

[1] The code for the simulation is available on *Github* (under an MIT license): `http://github.com/rvangijlswijk/ionospheric-ray-tracer`

Ray-tracing belongs to the domain of geometric optics and as such the rays adhere to the rules of optics. These rules follow from Fermat's principle (e.g. Mahajan 1998), and can be consolidated as follows:

1. Rays in free space travel in a straight line
2. Rays do not interfere with each other
3. Rays are reflected and refracted by objects that lie in their path
4. Ray paths between a transmitter and a receiver are reciprocal

Rays are originated from a *source* and end at a *sink*. In this numerical simulation, the beacons serve as a source by transmitting rays, and receiving antennas on mobile stations such rovers act as a sink. The path a ray travels in a scene is depicted in figure 5.1. In this figure, multiple objects are placed which interact with the ray. Collisions between the ray and objects are subject to Snell's law, causing the ray to reflect or refract. Whether refraction or reflection occurs depends on the index of refraction and the incident angle between the ray and the normal to the surface of the object. The index of refraction is defined primarily by the substance the ray is travelling through. Furthermore, the frequency of the ray influences the index of refraction. Rays with different frequencies will refract differently in the same substance, a phenomenon called *dispersion*.

When a ray transitions to a substance with a lower index of refraction, an incident angle, or *critical refraction angle* exists, below which no refraction occurs. This phenomenon is referred to as total internal reflection and causes the ray to fully reflect.

5.1.2 Algorithm

A software algorithm to trace the rays from source to sink is implemented using a Whitted-style ray-tracing algorithm as initially described by Whitted (1980). Rays are traced in a recursive manner until some boundary condition is met. A boundary condition for the purpose of this simulation can be that the ray collided with the terrain or disappeared into outer space. A pseudocode implementation of this algorithm is given in codeblock 5.1.

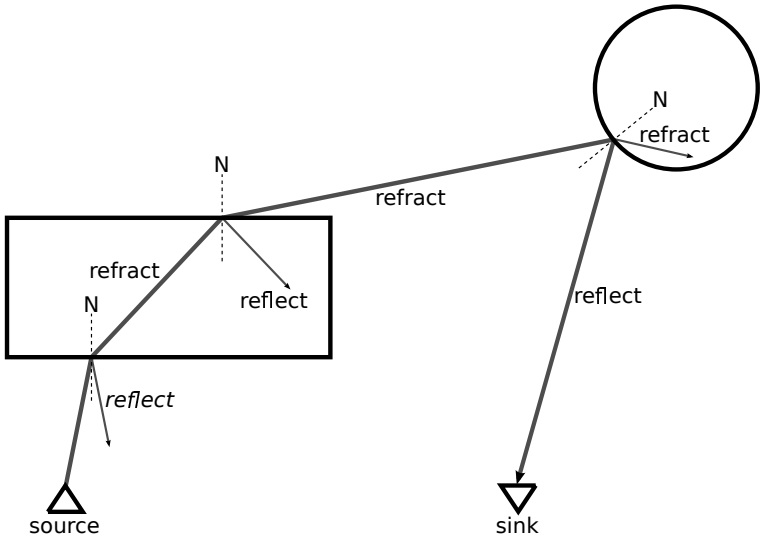


Figure 5.1: A schematic representation of the propagation of a ray in a scene containing multiple objects that interact with the ray.

```
1            void function trace(ray)
2                if ray is out of bounds
3                    stop;
4                intersection = scene.intersect(ray);
5                if intersection is ionosphere
6                    ionosphere.interact(ray);
7                    trace(ray);
8                if intersection is atmosphere
9                    atmosphere.interact(ray);
10                   trace(ray);
11               if intersection is terrain
12                   stop;
```

Codeblock 5.1: Implementation of Whitted-style ray-tracing algorithm

The most important feature in a Whitted-style ray-tracing algorithm for this numerical simulation is the recursive tracking of each ray. This allows for an implementation requiring little code.

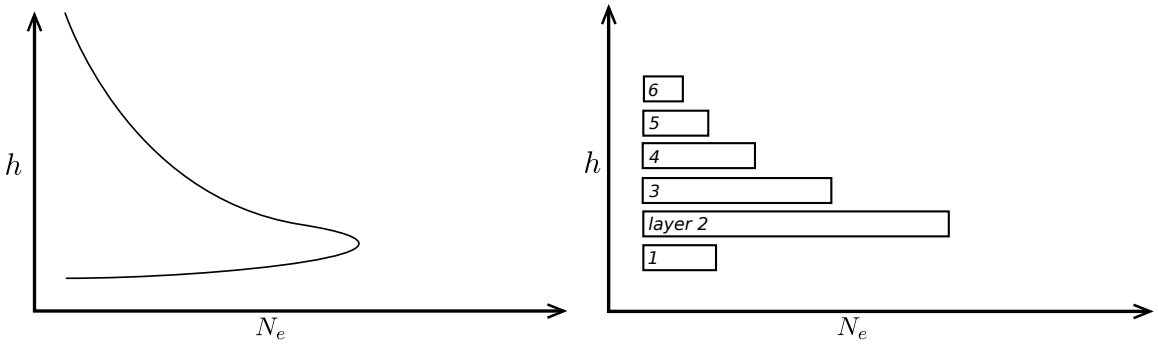
The interaction method shown in codeblock 5.1 determines how a specific scene object interacts with the ray. In this method, the interactions between the ray and its environment are numerically computed and applied onto the ray. Ultimately, the resultant ray is a sum of all the interactions of the ray with the objects in the scene.

5.1.3 Lamination

The ionosphere can be divided into a finite number of intermediate layers, each with constant physical properties. This process is referred to as *lamination* in the work of Kelso (1964). The individual ionospheric layers are referred to as *layers*. In this method, the ionospheric properties of individual layers are computed independently, using a set of linear equations. The total change of a ray is found by combining the interactions of all layers onto the ray.

The lamination method consists of a couple of steps. First, ionospheric layers are generated, each with a fixed set of parameters. For a given altitude h, the electron number density N_e, plasma frequency ω_p and collision frequency ν can be calculated. Now, let the thickness of the layer be given, and all effects on radio signals follow from a set of linear equations, previously defined in chapter 4. The lamination approach is used to model a Chapman-profile, by giving individual layers a certain height and electron density. Figure 5.2b demonstrates this process.

The method of lamination is only an approximation of reality and therefore only holds for a limited number of cases. The most obvious limitation is the number of layers that is used to approximate the ionosphere. A low number of layers is computationally cheap, but yields only a coarse approximation. The ionosphere can be better approximated by increasing the number of layers, but this also increases the computational cost of simulating the ionosphere. As a fixed number of computations have to be executed for each layer, the computational cost increases linearly with the number of layers. Hence, an optimum must exist for which the lamination method yields a reasonable approximation of reality while using



(a) An idealized Chapman profile *(b) A (simplified) laminated Chapman profile*

Figure 5.2: An idealized Chapman profile and a laminated approximation thereof. The Chapman profile in figure 5.2b is generated using six layers each with a constant electron number density.

as few computations as possible. In the simulation, the precision of the approximation is determined by the thickness of each layer. Each layer is given a specific vertical thickness, which determines the vertical region for which the constant ionospheric properties of this layer hold. The number of layers can thus be increased by decreasing the thickness of each layer. Hence, the layer thickness is the prime factor in defining both the precision of the approximation and the computational cost.

The optimal balance between the precision of approximation and computational cost can be found by comparing the ionospheric model produced by the lamination method with real measurement data. This exercise can be done for various layer thickness, and the optimum is subsequently found at the highest layer thickness which still accurately corresponds with real measurement data.

5.2 Optical effects

The dispersive nature of the ionosphere causes rays to refract and reflect. The refraction and reflection of a ray depend on the refractive indices of the two layers in between which a radio signal is transitioning. The values of these indices in turn depend on the properties of the ionospheric layer and the ray frequency. This yields a refractive index μ_1 and μ_2 for the first and second layer respectively. For calculations in the subsequent discussion it does not matter which of the two refractive indices is greater, what counts is that the first layer is the layer from which the ray originates, and the second layer is the one where the ray transitions to.

Figure 5.3 defines the model of ray reflection and refraction in 2D space. The incoming ray is denoted by $\boldsymbol{r_i}$ and the reflected and refracted rays by $\boldsymbol{r_r}$ and $\boldsymbol{r_t}$ respectively. Important to this discussion is that these ray vectors are all normalized. The surface normal $\boldsymbol{N}$ is always normal to the side of the surface facing the incoming ray. Based on this model the values for the direction of the $\boldsymbol{r_r}$ and $\boldsymbol{r_t}$ vectors can be derived.

The subsequent text treats the calculation of both vectors, as well as the critical angle and the magnitude of the reflected ray with respect to the refracted ray. The latter can be evaluated using Fresnel's equations. As the entire simulation revolves around a certain geometrical configuration, a definition of the coordinate system is given first. The section

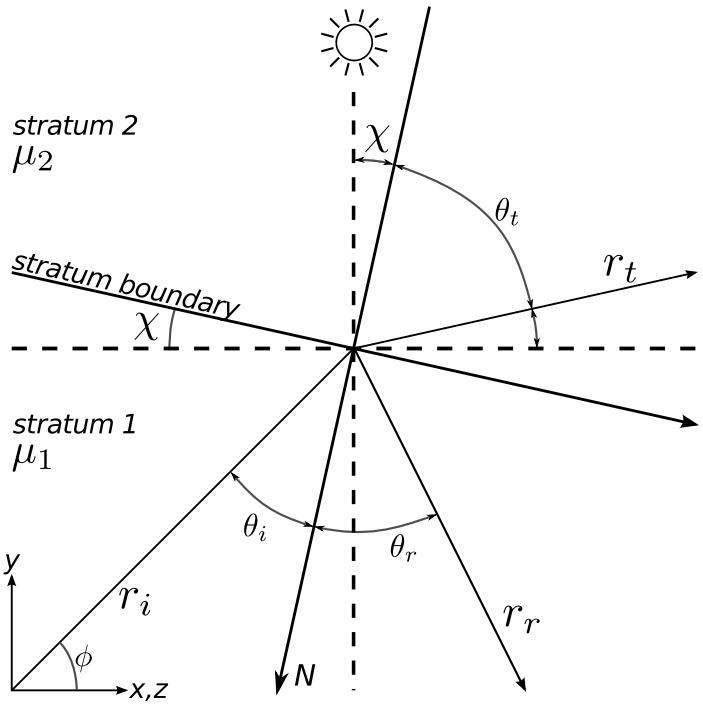


Figure 5.3: The refraction of a radio signal in an ionospheric layer in 2D space. The layer is modelled in the coordinate system described in section 5.2.1. The layer has an oblique angle w.r.t. the Sun vector. The SZA is indicated by χ, and the incident and refracted angles to the normal are denoted as θ_i and θ_r respectively.

concludes with a numerical analysis of the refractive indices, required to calculate both the direction of the refracted ray as well as the magnitude of both rays.

5.2.1 Coordinate system

A coordinate frame is selected for consistent calculations in the simulation, and such that these calculations can be made conveniently. The coordinate frame chosen is a Mars-centered ecliptic coordinate system, or MCECS. Its composition is depicted in figure 5.4. The main features of this coordinate system are its orientation with respect to the Sun and the Martian orbit plane. The xy-plane is coplanar with the orbital plane, whereas the origin of the coordinate frame is positioned in the center of Mars. The y-axis in this system is Sun-locked, pointing towards the Sun, and referred to as the Solar vector. The z-axis is orthagonal with respect to the orbital plane, pointing towards the ecliptic north pole. The x-axis completes the reference frame.

The angle which a mobile station has to the Solar vector is a main influence on the precision which can be obtained for position estimation. This angle is denoted as the Solar zenith angle SZA, or χ, and it influences the electron densities, as was explained in section 4.1.2. The electron densities in turn determine to what extend a radio signal is changed. Hence, it is convenient to select a vector pointing towards the Sun as a primary axis in the coordinate frame.

Furthermore, the position of a mobile station on the Martian surface can be described

by its latitude β and longitude λ with respect to the primary axes. The primary direction at 0° latitude equals the direction of the Solar vector. These angles, placed in the coordinate frame as described, subsequently provide a solid foundation to describe both the position and the estimated position of any mobile station on the Martian surface.

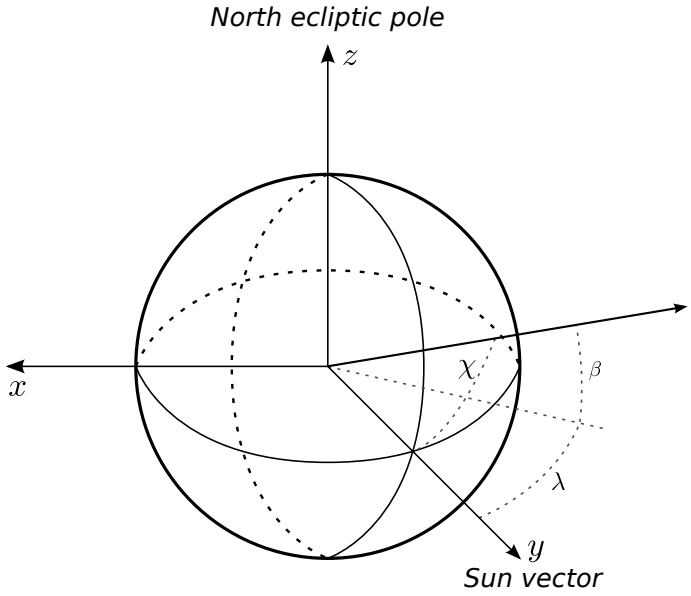


Figure 5.4: The Mars-centered ecliptic coordinate system used in the simulation. The y-axis is Sun-locked while the z-axis points to the ecliptic north. The x-axis completes the reference frame.

5.2.2 Reflection

The reflected ray is a function of the relationship between the incoming ray, the surface normal and the reflection angle (Glassner, 1989; Riedl, 2009). The angle of the reflected ray is equal in magnitude and opposite to the normal with respect to the incoming ray. Hence:

$$cos(\theta_r) = \mathbf{r_i} \cdot \mathbf{N} \tag{5.1}$$

A well-known property of reflection on a surface is that the angle of refraction equals the angle of incidence, or $\theta_i = \theta_r$. As the angles of incidence and reflection are equal, the parallel and perpendicular components of $\boldsymbol{r_i}$ and $\boldsymbol{r_r}$ with respect to the plane of incidence are equal in magnitude, but the perpendicular component of $\boldsymbol{r_r}$ has an opposite sign. This is denoted

$$\begin{aligned}
\mathbf{r_r} &= \mathbf{r_{i\parallel}} - \mathbf{r_{i\perp}} && (5.2)\\
&= (\mathbf{r_i} - (\mathbf{r_i} \cdot \mathbf{N})\mathbf{N}) - (\mathbf{r_i} \cdot \mathbf{N})\mathbf{N} && (5.3)\\
&= \mathbf{r_i} - 2(\mathbf{r_i} \cdot \mathbf{N})\mathbf{N} && (5.4)\\
&= \mathbf{r_i} - 2\cos(\theta_i)\mathbf{N} && (5.5)
\end{aligned}$$

5.2.3 Refraction

Unlike reflection, ray refraction is governed by the refractive indices of the two materials between which a ray is transferring. The angle of refraction is defined by Snell's law:

$$\frac{\sin(\theta_t)}{\sin(\theta_i)} = \frac{\mu_1}{\mu_2} \tag{5.6}$$

$$\sin(\theta_t) = \frac{\mu_1}{\mu_2}\sin(\theta_i) \tag{5.7}$$

Equation 5.1 is similar for the refractive case:

$$cos(\theta_t) = \mathbf{r_t} \cdot \mathbf{N} \tag{5.8}$$

The refracted vector $\boldsymbol{r_t}$ is then composed of:

$$\boldsymbol{r_t} = \boldsymbol{r_{t\parallel}} + \boldsymbol{r_{t\perp}} \tag{5.9}$$

$$\boldsymbol{r_{t\parallel}} = \frac{\mu_1}{\mu_2}|\ \boldsymbol{r_i}\ \parallel| \tag{5.10}$$

$$= \frac{\mu_1}{\mu_2}\left(\boldsymbol{r_i} + \cos(\theta_i)\boldsymbol{N}\right) \tag{5.11}$$

$$\boldsymbol{r_{t\perp}} = -\sqrt{1 - |\boldsymbol{r_{t\parallel}}|^2}\boldsymbol{N} \tag{5.12}$$

The parallel component in equation 5.12 can be replaced by making use of Snell's law from equation 5.7. Furthermore, taking the normalized parallel component as $\boldsymbol{r_{t\perp}} = \sin(\theta_t)$:

$$\boldsymbol{r_{t\perp}} = -\sqrt{1 - \sin^2(\theta_t)}\boldsymbol{N} \tag{5.13}$$

$$= -\sqrt{1 - \left(\frac{\mu_1}{\mu_2}\right)^2 \sin^2(\theta_i)}\boldsymbol{N} \tag{5.14}$$

$$\boldsymbol{r_{t\perp}} = -\sqrt{1 - \left(\frac{\mu_1}{\mu_2}\right)^2 (1 - \cos^2(\theta_i))}\boldsymbol{N} \tag{5.15}$$

$$\tag{5.16}$$

The solution for $\boldsymbol{r_t}$ comes from combining equations 5.9, 5.10 and 5.12:

$$\mathbf{r_t} = \frac{\mu_1}{\mu_2}\mathbf{r_i} + \left(\left(\frac{\mu_1}{\mu_2}\right)^2 \cos(\theta_i) - \sqrt{1 - \left(\frac{\mu_1}{\mu_2}\right)^2 (1 - \cos(\theta_i)^2)}\right)\mathbf{N} \tag{5.17}$$

5.2.4 Critical reflection angle

A noticeable limitation stems from Snell's law for certain angles of refraction. For angles of incidence where the sinus of the angle of refraction is greater than unity, Snell's law cannot be satisfied, as is evident from equation 5.7. This situation can occur when a ray travels to an ionospheric layer with a lower index of refraction. The effect is referred to as *total internal reflection*. An angle exists beyond which this effect occurs. This angle is referred to as the

critical reflection angle, and it is the angle at which $\sin(\theta_t) = 1$, or $\theta_t = 90°$. It can thus be derived from Snell's law:

$$\theta_{cr} = \sin^{-1}\left(\frac{\mu_2}{\mu_1}\sin(\theta_i)\right) \tag{5.18}$$

$$= \sin^{-1}\left(\frac{\mu_2}{\mu_1}\right) \tag{5.19}$$

Rays which intersect with an ionospheric layer greater than the critical reflection angle given in equation 5.19 will be reflected instead.

5.2.5 Fresnel equations

In any transition from one ionospheric layer to the next, only a fraction of the energy of a ray is transmitted whereas the rest is reflected. The Fresnel equations establish the ratio of transmitted (T) to reflected energy (R) on an interface with changing refractive index. The ratio follows from the conservation of energy:

$$T = 1 - R \tag{5.20}$$

The Fresnel equations account for a polarized electromagnetic signal with parallel and perpendicular components. For the reflectance, the fraction of these components in relation to the incoming energy are given as follows (Woan, 2000):

$$r_{\parallel} = \left(\frac{\mu_1 cos(\theta_i) - \mu_2 cos(\theta_t)}{\mu_1 cos(\theta_i) + \mu_2 cos(\theta_t)}\right)^2 \tag{5.21}$$

$$r_{\perp} = \left(\frac{\mu_2 cos(\theta_i) - \mu_1 cos(\theta_t)}{\mu_2 cos(\theta_i) + \mu_1 cos(\theta_t)}\right)^2 \tag{5.22}$$

In this simulation it is assumed that the rays are unpolarized, as the effects of taking into account polarization are far too small to be significant. As a result, the reflectance fraction is simply the average of its components:

$$R = \frac{r_{\parallel} + r_{\perp}}{2} \tag{5.23}$$

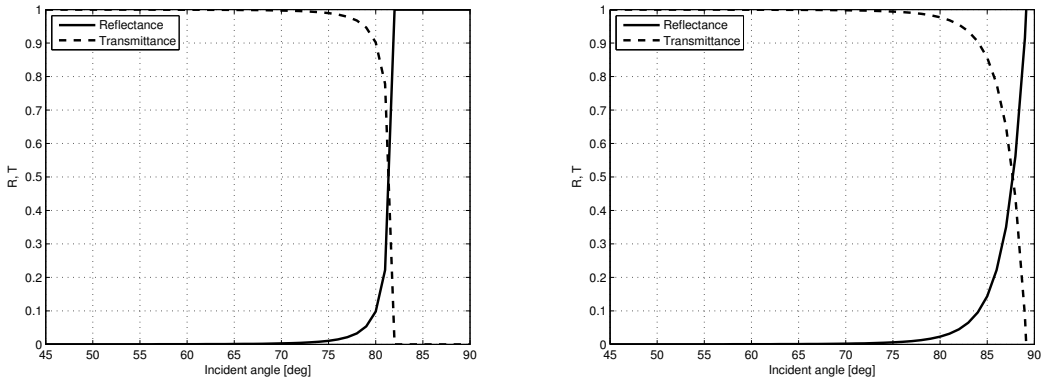
Equation 5.23 follows the limitations imposed by the effect of total internal reflection, see section 5.2.4. The critical reflection angle denotes the incident angle above which all energy is reflected. In this case, $R = 1$ and consequently $T = 0$.

Equations 5.20 to 5.22 describe the ratios of reflected to transmitted energy. Figure 5.5 depicts the ratios for the two end cases where a ray transitions to a medium with a lower and a higher index of refraction respectively. The former occurs when when a ray is departing from the peak density, whereas the latter occurs a ray is approaching the peak density of an ionospheric layer.

The Fresnel equations help the simulator validate whether to reflect or transmit, or both. For angles of incidence less than $75°$, generally transmittance can be assumed, as can be seen in figure 5.5. Above these angles the situation changes. Clearly, rays propagating almost straight up will mostly be reflected.

Figure 5.5 gives rise to an effect which can be exploited. In both cases $\mu_1 > \mu_2$ and $\mu_2 > \mu_1$, the reflectance is in some cases insignificant with respect to the transmittance.

Therefore, in the simulation an assumption can be made that only transmittance occurs without loss on precision. Such an assumption can be made when $T >> R$. As can be seen from the figure, this only holds for incident angles generally below $75°$, in which cases the reflection can be completely ignored.



(a) $\mu_1 > \mu_2$. *(b)* $\mu_2 > \mu_1$.

Figure 5.5: The governing behavior for the two cases where a ray transitions to an ionospheric layer with a lower refractive index (figure 5.5a, $\mu_1 > \mu_2$*) and a higher refractive index (figure 5.5b,* $\mu_2 > \mu_1$*) respectively. In figure 5.5a, total internal reflection occurs beyond* $\theta_i \approx 82°$*. The angles are measured with respect to the normal of the ionosphere, according to figure 5.3.*

5.3 Monte-Carlo approach

A Monte-Carlo method is applied to simulate the propagation of rays through the Martian environment which is exposed to a large number of disturbances. A large number of rays is initialized, each with a distinctive initial SZA and frequency. These rays are all fed to the simulator and the results are aggregated to yield a Gaussian distribution for a certain output parameter.

An example of this method is depicted in figure 5.6. In this figure a ray r is generated at a transmitter and received at the receiver as r'. The output metric is the distance d. Rays are generated with a certain launch elevation ϕ and frequency f. These are each subject to ionospheric fluctuations and geometrical path distortions. For example, the ionospheric composition can change in between the generation of any two rays. Furthermore, rays with a different launch elevation can possibly have the same distance, depending on which launch elevation is chosen. It is assumed that the dispersion of the resulting ray at the receiver r' is Gaussian distributed, as the signals are subject to random noise. This is a typical assumption, which follows from the *central limit theorem* (Grinstead and Snell, 2012). The central limit theorem states, that a large number of mutually independent random variables are well-approximated by a Gaussian function. The random noise can come from various sources. The ionospheric electron density varies continuously varies a little, and thus induces some random noise. Fluctuations in the collision frequency are another source of noise, as is the electron peak altitude.

After the ray tracing simulation has finished, the results are aggregated, resulting in a

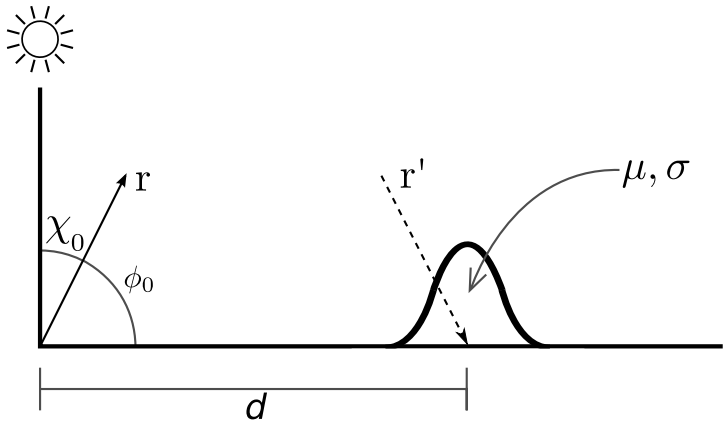


Figure 5.6: Depiction of a Monte-Carlo approach to ionospheric raytracing. A ray r is launched from a transmitter, placed at the subsolar point, with SZA χ_0 and launch elevation ϕ. Multiple rays are dispersed on the surface according to a gaussian distribution with mean μ and standard deviation σ at distance d from the origin.

Gaussian distribution of the estimated distance with a mean μ and standard deviation σ. This information is used to compute a model fit of the ray parameters over a metric of interest, which can be the distance of the mobile station to the beacon, for example. The model fit subsequently allows for derivation of the precision of position estimation, an invaluable metric used for testing the hypothesis.

The advantage of a Monte-Carlo approach is that results can be generated with repeated random sampling of rays. As the number of unknowns and disturbances are very high in a ray propagation model, it is hard to approach the problem analytically. The Monte-Carlo approach, however, allows the addition of many independent random factors, yielding a set of result samples without having to express an analytical model of the propagation path itself.

5.4 Simulation architecture

The task of the simulation is to trace rays through the Martian environment and calculate the way they change. This feat is achieved by splitting up the software in a number of modules, each with a specific responsibility. The major modules are the application base framework, the tracing module, the geometry module and the scene module. These are further elaborated in this section. An overview of these modules and their relations is given in figure 5.7. The section concludes with some notes on loss of precision, an effect which can greatly negatively influence the precision of the simulation results.

Figure 5.7 depicts the life of a ray as it is processed by the simulation. Rays are created by the application framework, according to the input parameters provided by the operator. Each ray starts at a certain position in the scene, and given a predefined direction. The scene module executes the trace method of all rays, to detect with what objects the rays intersect and interact. Once a collision is detected, the ray is passed on to the geometry module, which computes how the ray interacts with the object it collided into. The geometry module does so by calling the interaction method of the relevant object type; an ionospheric

layer for example. This may cause the ray direction to change due to refraction or reflection. Once interaction is done the ray executes its trace method again. This procedure is continued until all rays have collided with the terrain or a receiving antenna, at which point the simulation is halted.

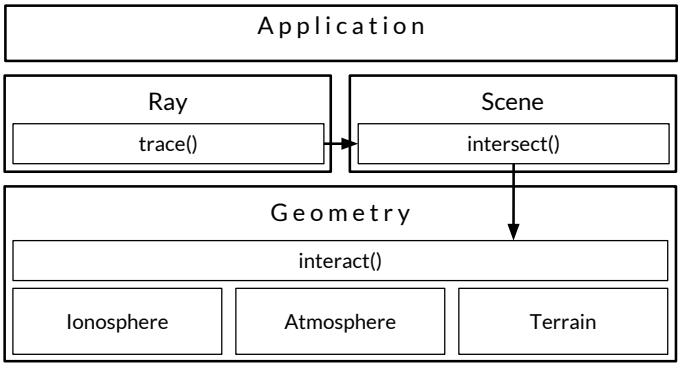


Figure 5.7: An overview of the software architecture.

5.4.1 Application framework

The application framework provides general support for running the simulation. As such, it has several responsibilities. First, it creates a *scene* which contains all the environmental features which are of relevance in the simulation. Examples of these features are ionospheric and atmospheric layers, as well as the terrain. The composition of these environmental features are based on a *scenario*. A scenario describes the composition of the environment. For example, it allows setting the ionospheric peak density, the solar longitude, atmospheric pressure and other properties of relevance. A *SceneManager* is used to keep track of all the environmental features and their location.

Once the scene is populated, the framework creates a number of ray objects with predefined physical parameters. These are for example frequency, signal strength and direction. The operator of the simulation can set the number of rays and their properties as he desires. There are two ways of defining ray objects; a direct method and a radiometric method. The direct method involves defining independent ray objects with independent properties. The radiometric method is geared towards simulation of antenna radiation patterns. In this method, an antenna object is defined with certain physical properties. Furthermore, it is given a certain radiation pattern. This radiation patter is then translated into a finite number of rays, each representing the radiated radio signal in a specific direction.

Once all the rays are created, they are each assigned to a *worker*. The workers take a ray as an input and calculate its entire propagation from creation to collision with the surface. This allows that the simulation can be run in a multi-threaded fashion, by having multiple workers run in parallel.

5.4.2 Tracing module

The main responsibility of the tracing module is to perform the ray-tracing algorithm, described in section 5.1.2, for radio signals, each represented as a single ray object. Their physical properties are encapsulated within the rays. Once a collision with an object in the scene

is detected, the tracing module lets that object modify the physical properties of the ray. This process continues until the ray either collides with the terrain or disappears in outer space. At this point the tracing algorithm for this ray is stopped.

In order to detect with which object a ray collides, the tracing module works together closely with the scene module. The tracing module informs the scene module that a ray is traveling from a certain point in a certain direction. The scene module performs an intersection calculation onto all geometry objects in the scene. Next, the tracing module passes the ray object to the interaction method of this geometry object to compute the effects it has on the ray and update the ray accordingly.

5.4.3 Scene module

The scene module keeps track of all environmental features in the scene and computes collisions between rays and geometry objects. For each ray, this module computes a list of geometry objects that lie in its path and possibly collide with the ray. Subsequently the geometry object closest to the ray origin is selected and identified as the object with which the ray collides. As soon as an object is identified, it is passed on to the tracing module such that the interaction method can be performed.

5.4.4 Geometry module

The geometry module defines information for each type of environmental feature and how these affect the physical properties of a ray. Each of these features is described by a *geometry* object. This object contains a model of the feature, how it affects a ray and its position in the scene space. The tracing module uses the geometry object to compute the effects it has on a ray object.

A geometry object must be defined for every type of object in the scene. Predefined example types are the atmosphere and the ionosphere. Each object type implements an interaction method, which receives a ray and modifies it according to the responsibilities of that object. For example, an ionospheric object attenuates and refracts the ray. The software allows for expansion of whatever object types may be required in a scene; a new object type only needs to implement the interaction method.

5.4.5 Computational precision

The precision of the calculations made in the simulation is influenced by the way the individual variables are stored. Variables are stored in datatypes, which each have a specific precision. Using the wrong datatype can cause a loss of precision. This naturally results in loss of precision in the simulated model. Therefore, workarounds need to be provided for variables which may encounter a loss of precision under certain conditions or edge cases.

The aforementioned is a problem in particular for the refractive index, as its values approach zero in some circumstances. In this case, a very high precision is required, which standard datatypes fail to provide. A failsafe workaround is to reduce the number of decimals used to store the refractive index.

$$\mu = N \cdot 10^{-6} \tag{5.24}$$

The substitute refractive index described in equation 5.24 avoids the loss of precision by breaking down the refractive index in separate parts and storing the fraction separately. Note that this approach only holds for $\mu < 10^{-6}$.

5.4.6 Verification

The calculations in the simulation are verified with the use of *unittests*[2]. Unittests are verification tests, which test each calculation method in the simulation individually. For example, one unittest can assert if the calculation of the plasma frequency in the simulation is done correctly.

The development approach, used for the development of the simulation, is *Test-driven development*. Without going too much into detail, this approach entails that the development of calculation methods is based on developing verification tests for them. The procedure for developing a calculation method, is to first define the testing code. The testing code defines what output should be received, given a certain input. This testing code then calls the respective method and verifies if the output of the method matches the expected output. Development of the calculation methods is generally done after defining the testing code. This way, one ensures up-front that the calculations are understood and that the method can be verified.

6 — Ionospheric Simulation Results

A validation of the simulation is necessary to assure that the produced data corresponds to existing research results. Once validated, the initial results from the simulator are used to explore boundary cases and limitations of the ionosphere as a radio signal carrier. As it turns out, a mobile station cannot always receive rays from beacons close by. Furthermore, the operational frequency is expected to impact both the path and absorption of a ray as it penetrates the ionosphere. These effects limit the operation of a positioning system.

In this chapter, the results obtained by the ionospheric simulation are discussed and validated. These results strongly depend on the nature of the ionosphere itself. Therefore, in section 6.1, several scenarios are defined which aim at expressing the nominal and extreme conditions in which a positioning system has to operate. A comparison of the simulated electron density and measured electron densities is performed in section 6.2. In section 6.3, the propagation path of the rays themselves is discussed. Here, the limitations and drawbacks of the operational frequency are explored. Finally, the absorptive levels, experienced by rays, are studied in section 6.4. Comparisons are made with the absorption levels as calculated in relevant research to validate the simulation.

6.1 Scenarios

The state of the environment severely impacts the propagation behavior of rays in the ionosphere, and hence, simulation results must be determined for various environmental compositions. These compositions are defined in *scenarios*. Table 6.1 gives an overview of the scenarios that have been defined, whereas this section discusses the necessity of each scenario.

A qualitative discussion about the Martian environment was given in chapter 3. Therein it was found that the following factors contribute significantly to the propagation of radio signals, and thus to the precision of position estimation:

- Fluctuations in electron peak density and altitude
- The presence of dust storms
- The time of day, influencing ionospheric composition
- The occurrence of Solar flares

Fluctuations in electron peak density and altitude are caused by both seasonal variations (Morgan et al., 2008) and the 11-year Solar cycle (Mendillo et al., 2004). It is imperative to study the operation of a positioning system both extreme cases of minimum and maximum Solar interaction, and therefore scenarios have been created to reflect on these extremes. Mars is at aphelion in its orbit during the Northern summer solstice, and at perihelion during the Southern summer solstice (Morgan et al., 2008, fig. 7g). During a low Solar activity, the peak density at perihelion is around $1.8 \cdot 10^{11} m^{-3}$, whereas at aphelion the peak density is lower, with $N_e = 1.4 \cdot 10^{11} m^{-3}$ (Morgan et al., 2008). The opposite effect is seen for the electron peak altitude, which increases from $h_0 = 115km$ to $h_0 = 134km$ at the transition from aphelion to perihelion (Morgan et al., 2008). The neutral scale height stays nearly constant during the entire Martian year. At times of Solar maximum, the peak

Table 6.1: Environmental parameters for all scenarios

Scenario	Layer	Electron number density	Electron peak altitude	Reference
		N_e	$h_{max,0}$	
Basic	M2	$2.5 \cdot 10^{11} m^{-3}$	$125km$	Williams (2015)
Nominal	M2	$1.8 \cdot 10^{11} m^{-3}$	$130km$	Williams (2015)
	M1	$9.0 \cdot 10^{10} m^{-3}$	$110km$	Morgan et al. (2008)
Max	M2	$2.6 \cdot 10^{11} m^{-3}$	$134km$	Mendillo et al. (2004)
	M1	$1.0 \cdot 10^{11} m^{-3}$	$110km$	Morgan et al. (2008)
Min	M2	$1.4 \cdot 10^{11} m^{-3}$	$115km$	Mendillo et al. (2004)
	M1	$9.0 \cdot 10^{10} m^{-3}$	$100km$	Morgan et al. (2008)
Dust storm	M2	$2.0 \cdot 10^{11} m^{-3}$	$137km$	Wang and Nielsen (2003)
	M1	$9 \cdot 10^{10} m^{-3}$	$115km$	
SEP event	M2	$1.75 \cdot 10^{11} m^{-3}$	$135km$	Morgan et al. (2008)
	M1	$1.2 \cdot 10^{11} m^{-3}$	$105km$	Mendillo (2006)

density at perihelion increases to values of $N_e = 2.6 \cdot 10^{11} m^{-3}$ (Mendillo et al., 2004). The compositions under a Solar minimum and Solar maximum are indicated with the *min* and *max* scenarios respectively.

Additionally, a nominal scenario has been defined to test the operation of a positioning system under nominal conditions. These conditions are based on the research by Safaeinili et al. (2007).

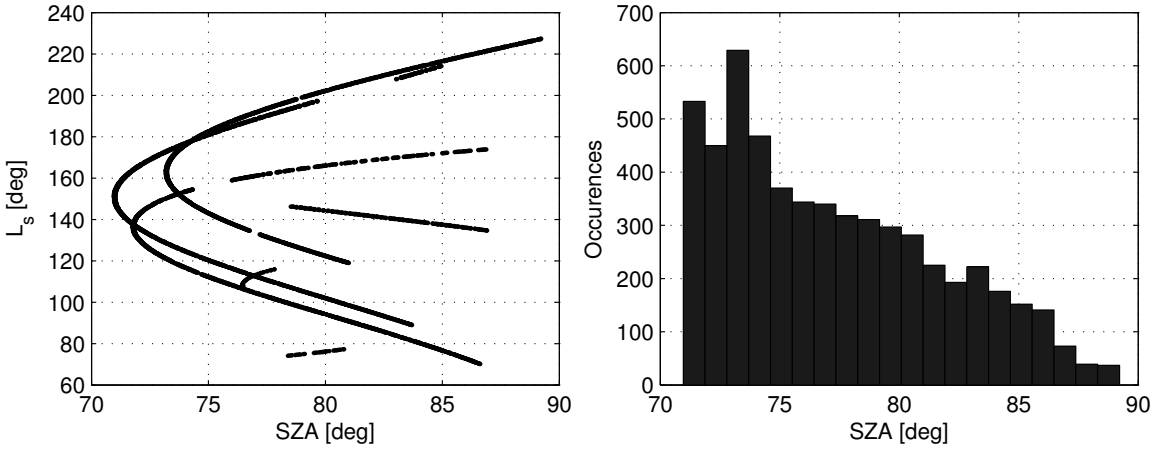
Dust storms have a profound effect on the composition of the ionosphere. Mainly, the electron peak altitude increases in magnitude during a dust storm (Wang and Nielsen, 2003). Data from a global dust storm in 1971-1972, recorded by Mariner 9, has shown that the electron peak altitude at the climax of the dust storm was as high as $137km$ (Wang and Nielsen, 2003) during that time of the Martian year. The electron peak density, however, does not change much during a dust storm. The values chosen for this dust storm scenario mimic the conditions of the ionosphere at the climax of the dust storm season reported by Mariner 9.

SEP events occur infrequently, causing photoionization due to EUV radiation and x-ray photons. The enhanced Solar flux during a SEP event mainly influences the the electron peak density of the M1 and EP layer (Mendillo, 2006). As the x-ray photons are highly energetic, they penetrate deep into the ionosphere, therefore mainly affecting the EP layer. The M1 layer is mainly influenced by EUV radiation, causing increases by as much as 200% in electron peak density (Mahajan et al., 2009). For the SEP event scenario, the data from the effects of the Solar flare on 15 April, 2001 is used. In this event, values were reported for the M1 layer corresponding to an electron peak density of $N_e = 1.2 \cdot 10^{10} m^{-3}$ at $\chi = 0$. The M2 layer stayed constant at a peak density of $N_e = 1.7 \cdot 10^{11} m^{-3}$.

The nightside of the red planet receives little to no Solar radiation at all. It is generally agreed that the nightside does not follow a Chapman profile, but is rather linear (Withers, 2011). Nevertheless, much is still unknown about the exact composition of the nightside. Therefore, this work will not take the nightside into account. It is recommended for futher work to expand the scenarios with a nightside scenario.

6.2 Electron density profiles

The electron density profiles generated by the simulation coincide with measurement datasets obtained from Mars missions[1]. A comparison with the electron density profiles obtained by the *Mars Global Surveyor* (MGS) mission is done to demonstrate the validity of the simulation. The MGS dataset only contains data for very large SZA angles, due to the nature of the radio occultation experiment that was used to obtain the data. Nevertheless, for these high SZA angles, many profiles are available. Figure 6.1 gives an overview of the situations for which the MGS dataset provides electron density profiles.



(a) Availability of electron density profiles over the course of the Martian year. The time is indicated by the solar longitude L_s, the angle of Mars in its orbit around the Sun.

(b) A histogram of data availability as a function of SZA . The angles are binned per degree. This histogram shows that the MGS dataset only has profiles available for $SZA > 70°$.

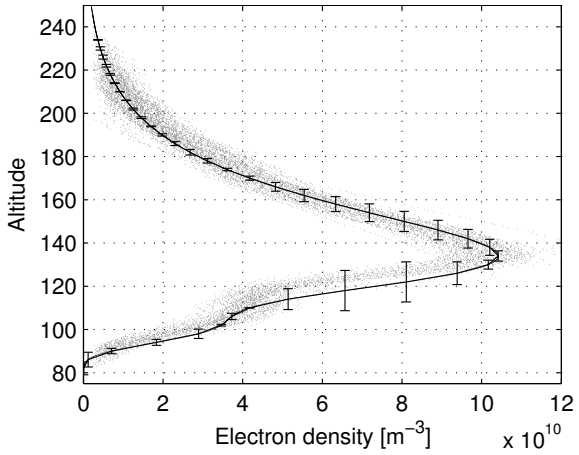
Figure 6.1: Availability and usability of the MGS dataset[1].

A comparison with the MGS dataset and an electron density profile generated by the simulation is shown in figure 6.2a for an $SZA = 71°$. An ionospheric composition consisting of an M1 and M2 layer is used to obtain the simulated profile. To obtain this profile, an electron peak density of $N_{e,max} = 1.7 \cdot 10^{11} m^{-3}$ is used for the M2 layer and $N_{e,max} = 6.0 \cdot 10^{10} m^{-3}$ for the M1 layer. Simultaneously, the M2 and M1 layer are given an electron peak altitude of $h_{max} = 121km$ and $h_{max} = 95km$ respectively. The secondary peak of the M1 layer is visible below the main peak both in the MGS dataset and the simulated profile. The simulated profile relates to the dataset with a mean standard deviation of $\bar{\sigma} = 2.19 \cdot 10^9 m^{-3}$. The highest discrepancy is found in the transition from the M1 to the M2 layer. At other altitudes, the standard deviation is significantly lower.

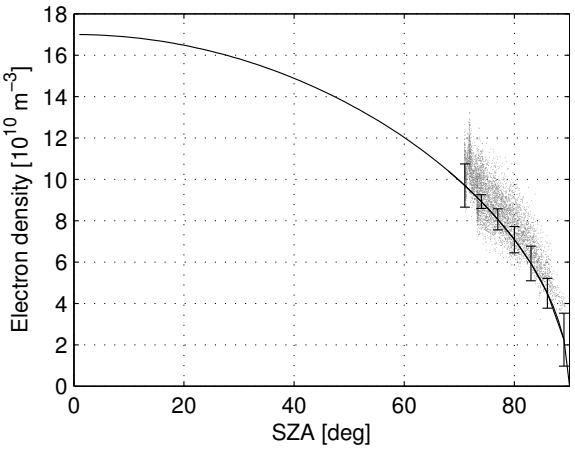
Figure 6.2b sets out the electron peak density as a function of the SZA , as found at the peak altitude. The density is expected to decrease with increasing SZA . This behavior follows general Chapman theory, previously described in section 4.1.2, and is also apparent from the MGS dataset shown in figure 6.2b. The simulated profile follows the curve of the dataset for angles of $SZA \geq 70°$. As mentioned before, this is due to the availability of the dataset. A mean standard deviation of $\bar{\sigma} = 8.13 \cdot 10^9 m^{-3}$ has been found. Due to the high resemblance between the simulated profile and the measurements obtained from the dataset, it can be assumed that the simulated profile is realistic for lower Solar zenith

[1]Data obtained from the Mars Orbital Data Explorer at `http://ode.rsl.wustl.edu/mars/index.aspx`

angles.



(a) Electron peak density profile for SZA$= 71^\circ$. $\bar{\sigma} = 2.19 \cdot 10^9 m^{-3}$.



(b) Electron density as a function of SZA. $\bar{\sigma} = 8.13 \cdot 10^9 m^{-3}$.

Figure 6.2: A comparison of an electron density profile obtained through MGS data and the simulated profile. Gray dots indicate measurements from the MGS dataset[2] whereas the black line depicts the simulated profile. Error bars indicate the deviation of the profile from the dataset.

6.3 Geometry

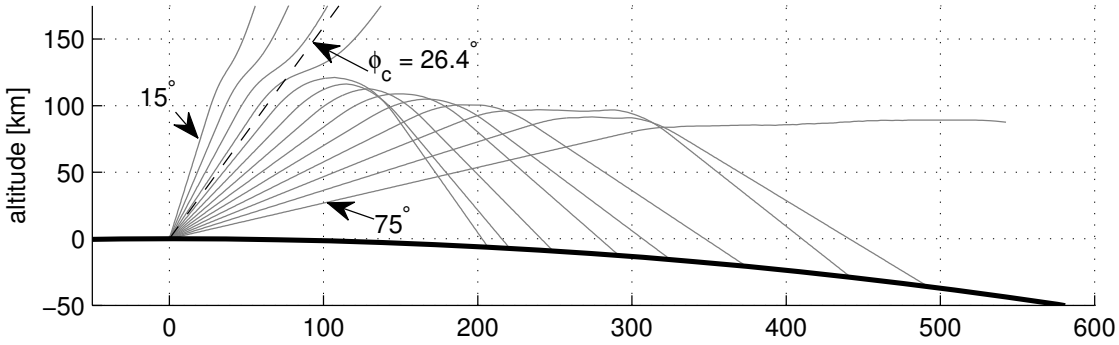
Several effects impose limitations on the usability of ray propagation. Two significant effects are the iris effect and the skip distance, the former of which has been treated in section 4.2.2. It is imperative to under which boundary conditions a localization system can operate. The iris effect, discussed in section 6.3.1, limits the elevation angles of rays which are successfully refracted to the surface. This results in a certain distance from a beacon where no rays are received. This effect is called the *skip distance* (Kelso, 1964), and it is discussed in section 6.3.2.

6.3.1 The iris effect

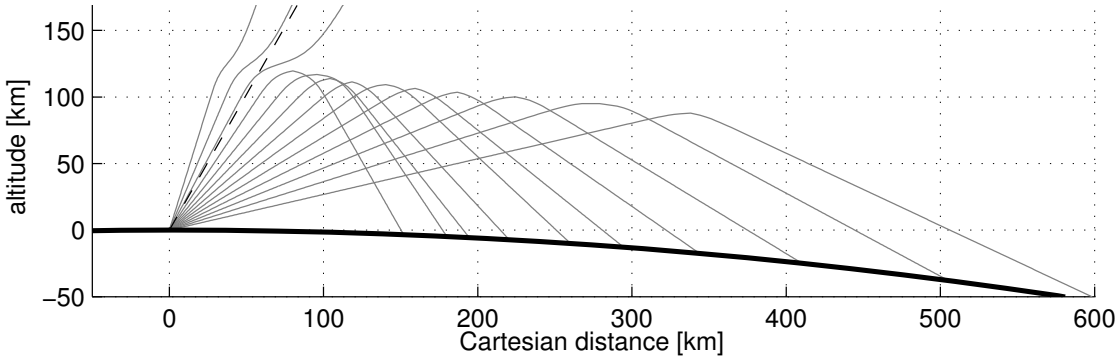
The simulated results for the iris effect follow the theoretical expectations, discussed in section 4.2.2. The effect is tested in the simulation, the results of which are shown in figure 6.3. This figure shows two ionospheric compositions; one containing both an M2 and M1 layer, and one containing only the major M2 layer. The plasma frequency for the M2 layer in both cases is $f_p = 4.49MHz$. Using a ray frequency of $f = 5.0MHz$, this results in a critical angle of $\phi_c = 26.4^\circ$ according to equation 4.29. The simulation results demonstrate equivalent results. Rays with a zenith angle of up to 25° penetrate through the ionosphere, whereas rays with a zenith angle of $> 25^\circ$ are refracted and hit the surface.

It was expected that the iris effect might not hold for ionospheres consisting of multiple layers, as each layer has a different plasma frequency at the layer maximum. However, it has been found that the iris effect is a robust figure. The results shown in figure 6.3a show that the iris effect still applies for a multi-layer composition. For both a single-layer and a multi-layer ionosphere, the critical angle found equals $\phi_c = 26.4^\circ$. The robustness follows

[2]Data obtained from the Mars Orbital Data Explorer at `http://ode.rsl.wustl.edu/mars/index.aspx`



(a) Nominal scenario with layers M2 and M1



(b) Basic scenario with only the M2 layer

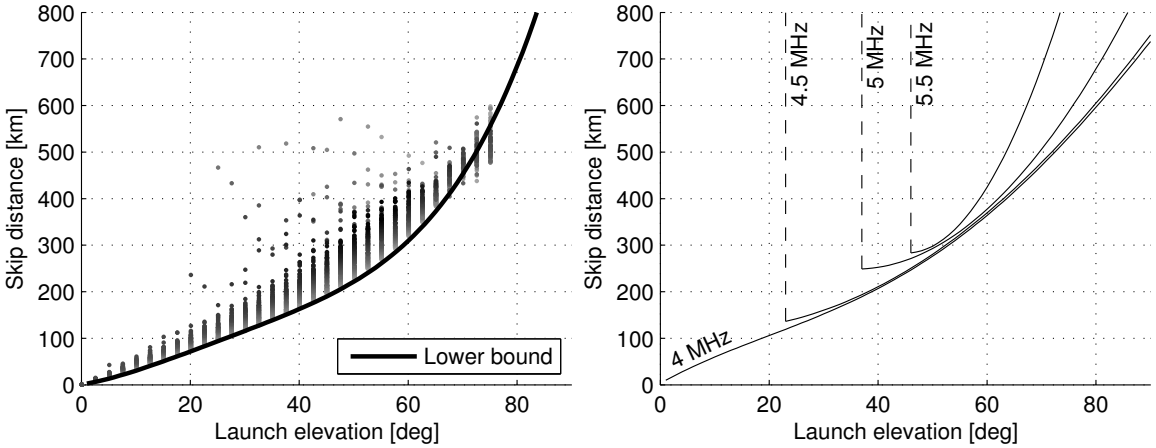
Figure 6.3: An illustration of the iris effect, with $\phi_c = 26.4°$. Plotted are 13 rays with a zenith angle from $15°$ to $75°$ in $5°$ intervals, originating from a beacon at coordinates $[0, 0]$. The rays are depicted by gray lines. For this plot, a ray frequency of $f = 5MHz$ is used, with an ionospheric composition consisting of the M1 and M2 layer under nominal conditions (see table 6.1). In figure 6.3a, the ray with a zenith angle of $75°$ is seen trapped at the layer maximum. Note, that the launch elevation and zenith angle are complementary.

from the fact that the M2 layer is really the dominating layer in the Martian ionosphere, and other layers have little effect on the critical angle, causing insignificant deviations.

6.3.2 Skip distance

The skip distance is a measure of the minimum distance to a beacon in which the mobile station cannot receive any signal from that beacon. This distance, indicated by d_s, increases with frequency. The skip distance is an indirect result from the iris effect, described in section 6.3.1, as some rays with a too steep launch elevation and a high frequency are not reflected by the ionosphere. This creates a circle around the beacon within which no rays can be received by a mobile station. The area within this circle is referred to as the *skip zone*.

The effects of the skip distance for a nominal scenario are illustrated in figure 6.4. The results are generated for ray frequencies ranging from $3MHz$ to $8MHz$. As can be seen, the skip distance increases with launch elevation. Furthermore, an increase in frequency causes an increase in skip distance. The gradient in figure 6.4a, which has darker tones for higher frequencies, illustrates this effect. Also, higher frequencies only yield results for



(a) A gradient of dots indicates the result of one ray with a certain ray frequency. Brighter tones indicate lower frequencies, darker tones indicate higher frequencies. A curve is fitted to the lower bound of the skip distance at $3.0MHz$.

(b) A model of the lower bounds and skip distances for select frequencies. A curve is fitted for each of these select frequencies, yielding a cutoff-launch elevation whose surface distance equals the skip distance.

Figure 6.4: Illustration of the skip distance, d_s, for various frequencies as a function of launch elevation. The results are gathered for a nominal scenario. The results are obtained for frequencies from $3MHz$ to $8MHz$, by using a second-order exponential curve fit, where $d_s = a\phi^2 + b\phi + c$. The plasma frequency for this scenario is $f_p = 4.45MHz$.

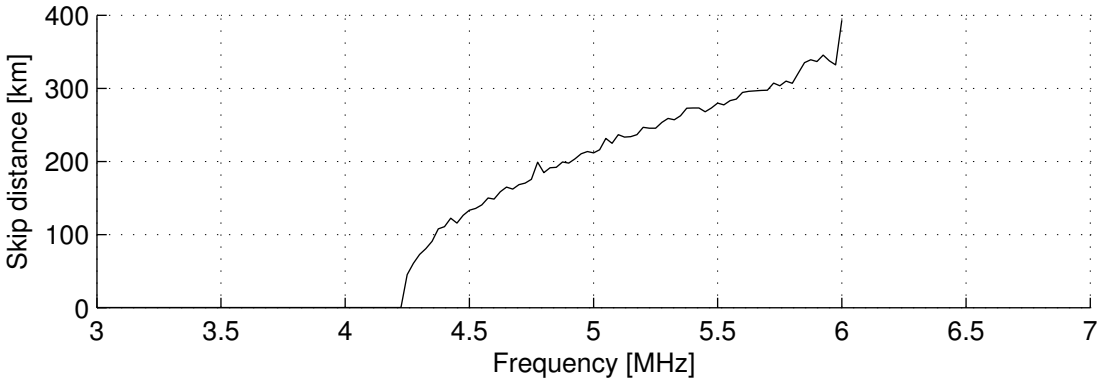


Figure 6.5: Illustration of the skip distance as a function of MUF . The plasma frequency is $f_p = 4.45MHz$. The skip distance increases for frequencies which are higher than the layer maximum plasma frequency.

higher launch elevations, due to the aforementioned iris effect. For each ray frequency, a lower bound exists below which no ray is reflected at any angle. This lower bound is indicated in figure 6.4b for select frequencies. This figure demonstrates in more detail that an increase in frequency will result in an increase of the lower bound.

A maximum frequency exists for each distance, at which a mobile station can still receive a signal from a beacon. Above this frequency, no signal will be received. This frequency is the *maximum usable frequency*, abbreviated MUF . Figure 6.4b helps to determine what maximum frequency can be used. The figure indicates, for several frequencies, the area within

which a signal can be received. For example, a distance of $d_s = 200km$ can only be achieved with frequencies of less than $4.5MHz$. Thus, the maximum usable frequency is $< 4.5MHz$. Clearly, the MUF is a convenient measure for selecting the operational frequency in a localization system.

Figure 6.5 illustrates the behavior caused by the MUF in more detail. Here, three regions with different behavior can be distinguished. First, there is the region below the plasma frequency. Rays with a frequency below the plasma frequency are refracted regardless of their incident angle w.r.t. the ionosphere. A ray is refracted at the layer maximum when its frequency equals the peak plasma frequency of the ionosphere. Hence, no skip distance exists for rays with frequencies below the plasma frequency.

The second region ranges from the plasma frequency to a frequency of $\approx 6MHz$. A parabolic increase is seen in this region. The change in behavior occurs close to the maximum plasma frequency, which is $4.45MHz$ at the layer maximum of the M2 layer for the nominal scenario. The skip distance is a most profound effect in this region. Positioning systems using system frequencies above the plasma frequency must thus take this effect into account.

Rays with a frequency above $6MHz$ do not appear in figure 6.4b at all. Their disappearance is caused by the fact that their frequency is too high to be refracted by the ionosphere. At such high frequencies, the rays penetrate the ionosphere completely, disappearing into outer space.

Ultimately, figure 6.4b demonstrates that a frequency region exists for which no skip distance exists. This knowledge is invaluable, considering the skip distance is a negative effect which should be reduced as much as possible. After all, mobile stations within the skip distance of a beacon will not receive its signals. Therefore, care should be taken when determining where to place beacons. Note that the results given in figure 6.4 and 6.5 apply specifically to the nominal scenario. Naturally the composition of the ionosphere influences the results. Therefore, different skip distances and maximum usable frequencies will be obtained for different scenarios.

6.4 Absorption

Results from the simulation show dissimilar results for absorption levels of reflected rays and non-refracted rays. The results for non-refracted rays have been verified with existing research. The results for reflected rays indicate that the main influences are frequency and surface distance.

6.4.1 Non-refracted rays

The absorption levels experienced by non-refracted rays are validated with previous research for various situations. Both ionospheric features, like the electron number density and scale height, and geometric features, like the Solar zenith angle, influence absorption levels significantly. Specifically, the work of Nielsen et al. (2007) and Withers (2011) is used for validation.

A ray experiences maximum absorption at the subsolar point in the ionosphere. For larger Solar zenith angles, the absorption decreases. The theory on absorption, described in section 4.3, showed the dependence of the absorption on the electron number density.

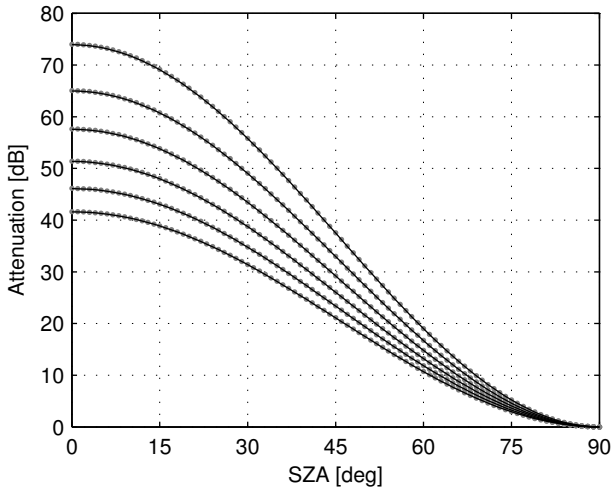


Figure 6.6: The simulated absorption as a function of the SZA compared with the results from Nielsen et al. (2007) for various frequencies. Black lines indicate simulated values, gray dots indicate results obtained by Nielsen et al. (2007). The ray frequencies used range from $3MHz$ to $4MHz$ with steps of $200kHz$. The lowest absorption is obtained with a frequency of $4MHz$. The environmental properties are $N_{e,max} = 1.65 \cdot 10^{11} m^{-3}$ and $H_n = 9.5km$ using an ionosphere with only a M2 layer. The launch elevation is $0°$.

This dependence enforces a relationship between the absorption and the Solar zenith angle, as the electron number density itself decreases with increasing SZA . These effects have been previously studied by Nielsen et al. (2007) and the results of his research are compared with the simulation results in figure 6.6. Clearly, the simulation results align closely with the results done by Nielsen et al., and thus it can be safely assumed that the simulation can accurately simulate the influence of the Solar zenith angle on the absorption. Figure 6.6 also demonstrates the impact of the ray frequency on absorption. An increase in frequency causes a decrease in absorption, and thus higher frequencies are beneficial for mitigating the absorption levels.

When dissecting the ionosphere into its individual layers, it is found that each layer has a different impact on absorption. Withers (2011) studied the impact of individual ionospheric layers on absorption levels, as experienced by the MARSIS instrument on-board the *Mars Express*. His research found that the M2 layer causes the lowest absorption, followed by the M1 and meteoric layer, which cause comparable absorption levels. However, the highest absorption is obtained by the EP layer, residing at an altitude of $35km$. The absorption measurements obtained from the simulation yield comparable results, as illustrated in figure 6.7. In this figure, absorption levels are simulated for various electron peak densities, ranging from $1 \cdot 10^7 m^{-3}$ to $1 \cdot 10^{12} m^{-3}$ on a logarithmic scale. It is expected that an increase in electron density causes an increase in absorption, and this is exactly what is seen in the results in figure 6.7.

Table 6.2 gives an overview of the statistical differences between the results obtained by Withers (2011) and those from the simulation. The differences between the two models are clearly small for both the M2 and M1 layer. However, a significant discrepancy is seen for the meteoric layer. This discrepancy can be explained by considering the different ap-

Table 6.2: Statistical results of the differences in absorption levels between the model provided by Withers (2011) and the simulation results. Deviations are given in m^{-3}. The differences are indicated by σ. A different scenario was used for each combination of N_e and ionospheric layer, to create the same conditions as in the research of Withers. In every case, only a single-layer ionosphere was simulated.

N_e	M2 layer		M1 layer		Meteoric layer		EP layer	
	σ	%	σ	%	σ	%	σ	%
$\mathbf{10^7}$	$7.00 \cdot 10^{-4}$	3.63%	$2.40 \cdot 10^{-3}$	0.64%	$2.23 \cdot 10^{-2}$	5.95%	$3.61 \cdot 10^{-2}$	3.82%
$\mathbf{10^8}$	$2.20 \cdot 10^{-3}$	0.76%	$7.70 \cdot 10^{-3}$	0.72%	$7.04 \cdot 10^{-2}$	5.94%	0.114	3.82%
$\mathbf{10^9}$	$6.90 \cdot 10^{-3}$	0.74%	$2.44 \cdot 10^{-2}$	0.71%	0.223	5.94%	0.361	3.82%
$\mathbf{10^{10}}$	$2.17 \cdot 10^{-2}$	0.72%	$7.72 \cdot 10^{-2}$	0.71%	0.704	5.94%	1.14	3.82%
$\mathbf{10^{11}}$	$2.38 \cdot 10^{-2}$	0.72%	$8.81 \cdot 10^{-2}$	0.72%	0.687	5.80%	2.33	4.66%
$\mathbf{10^{12}}$	$2.38 \cdot 10^{-2}$	0.72%	$8.85 \cdot 10^{-2}$	0.72%	0.675	5.78%	5.91	5.48%

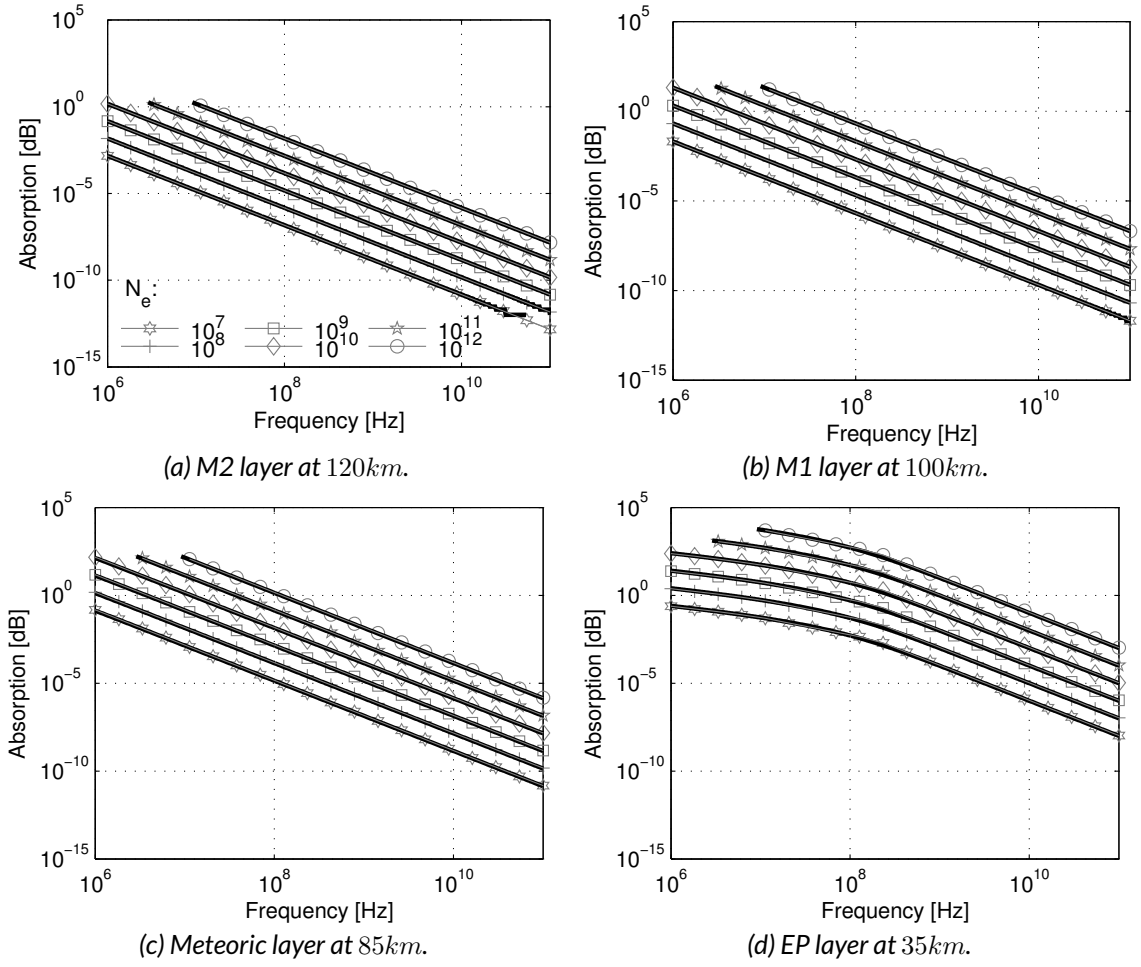
proach to calculating the collision frequency. The simulation makes use of the approach first described by Schunk and Nagy (2009) (see section 4.1.4). On the contrary, the approach for the collision frequency, adopted by Withers, uses a simplification by assuming a constant surface temperature. Furthermore, he assumes an exponential decrease in collision frequency with increasing altitude. This causes a difference in absorption results, especially at lower altitudes. At these altitudes the collision frequency has a greater magnitude and thus more severely impacts the resulting absorption calculation. The effects of these simplifications are seen in the comparison of the EP layer as well. In this layer, the standard deviation between the simulated and Withers' results is even higher.

6.4.2 Reflected rays

Reflected rays encounter more absorption with increasing frequency and surface distance. These findings are conflicting with the results found for non-reflected rays, but the theory behind their dissimilar behavior holds a common ground.

Figure 6.8 shows the relationship in a frequency spectrum between the absorption and the surface distance which a ray covers. Interestingly, the absorption increases with decreasing surface distance. Furthermore, the absorption is also seen to increase with increasing frequency. The latter is visible in the figure for rays whose surface distance is $200km$ or larger. In the lower end of the frequency spectrum, the rays are bounded by the plasma frequency. As was derived from equation 4.28 in section 4.2.1, rays cease to exist when the refractive index becomes imaginary. As a ray intersects the ionosphere at an oblique angle, one cannot simply compare frequency of the ray. Instead, the component of the wavelength parallel to the normal of the ionospheric surface is compared. This effect causes the lower bound in figure 6.8. The upper bound is a direct consequence of the skip distance. As the frequency increases, the skip distance increases. Thus, ray with a lower surface distance are not reflected and instead penetrate the ionosphere.

The correlation between absorption and frequency stems from the penetrating nature of radio signals. Rays with a low enough frequency are seen refracting off the ionosphere, as is demonstrated earlier in this work. If one now were to increase the ray frequency bit by bit, one would see that the signal penetrates the ionosphere deeper before being reflected. At



(a) M2 layer at $120km$.

(b) M1 layer at $100km$.

(c) Meteoric layer at $85km$.

(d) EP layer at $35km$.

Figure 6.7: A comparison of simulated absorption levels and those obtained from Withers (2011) for a wide range of frequencies. The results shown are for rays which penetrate the ionosphere. The absorptions for various electron peak densities obtained by Withers are displayed in gray. The scale height used is $H_n = 7.6km$, and the launch elevation is $0°$. The electron density, N_e, is given in m^{-3}.

some point, the signal has a high enough frequency to penetrate the entire ionosphere without being reflected. Because the ray penetrates the ionosphere deeper, it encounters more collisions, and thus it is more absorbed than a signal which was reflected at a lower altitude. Hence, reflected rays experience more absorption when their frequency is increased.

The relationship between absorption and surface distance is a result from the geometric construct which causes the surface distance in the first place. As was established, the surface distance of a signal follows from its initial launch elevation. The higher the launch elevation, the larger the surface distance is which the radio signal covers. Recall figure 6.3, which clearly shows this behavior. As with the frequency relation, a higher launch elevation will cause the ray to penetrate the ionosphere more deeply before being refracted. As a result, it is absorbed more severely.

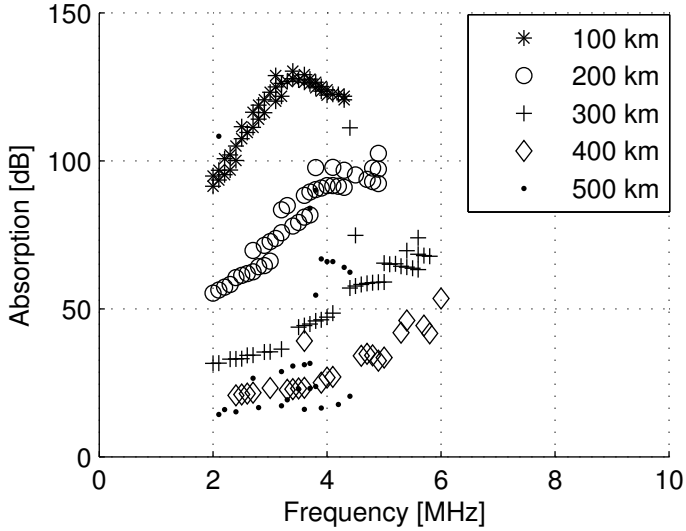


Figure 6.8: In a frequency-absorption spectrum, the absorption is seen to be increasing for both frequency and distance. Plotted are the total absorptions of rays as they hit the surface, grouped by surface distance they traveled. The reflected rays are lower bounded by the plasma frequency and upper bounded by the skip distance. The results are obtained with a simulation of the ionosphere under the nominal scenario.

6.5 Summary

The discussion in this chapter so far has shown that the choice of a frequency, which is ultimately optimal for use in a localization system, depends on the skip distance and the magnitude of absorption experienced. From the results above, it is clear that higher frequencies are not beneficial. For coverage purposes, one wants to decrease the frequency as much as possible as the skip distance only occurs at frequencies above the plasma frequency. Simultaneously, higher frequencies yield higher absorption levels. This in turn means that there is a higher probability that a mobile station cannot receive a signal from a certain beacon. Overcoming higher absorption levels ultimately means bigger and heavier mission hardware, something which is a severe drawback for hardware that is to be deployed on Mars. In subsequent chapters, the impact of the skip distance and operational frequency will be studied, so that an operational optimum can be found for a localization system on Mars.

7 — Antenna design

Position estimates can only be established when the mobile station is able to receive beacon signals. The ability to receive a signal depends on the received signal strength, a factor largely influenced by the antenna that is used to both broadcast and receive a signal.

In this chapter, an ideal antenna model is proposed which is used in the subsequent research. It is out of the scope of this research to treat the effects of antenna performance in detail. Hence, no detailed trade-off between antennae models will be performed. Instead, a simple ideal antenna will be selected, such that the implications and limitations of the use of an antenna can be demonstrated.

Section 7.1 aims at identifying the formulas that are important to calculate the performance of a given antenna. Subsequently, an antenna is selected for use in this research. Afterwards, the signal strength between the beacon and the mobile station is studied. The connection between a beacon and mobile station is referred to as a *link*. The link consists of gains and losses in signal strength, which together make up the *link budget*. Ultimately, the goal is to find the maximum *path loss* that still leads to a successful reception of a signal at the mobile station. Section 7.2 aims at finding this maximum path loss. The value found will be used in the research as a lower bound to discard results of signals which do not achieve the lower bound. Finally, a short summary of the findings on antenna design is given in section 7.3.

7.1 Antenna selection

Antennas radiate energy in certain directions. How this energy is radiated, depends on the type of antenna that is used. This section first covers the main equations that govern the radiation of energy, and which are relevant to this research, in subsection 7.1.1. Furthermore, an antenna is selected for the use in this work in subsection 7.1.2

7.1.1 Theory

The performance of an antenna is measured by the amount of energy which it radiates, and in what direction that energy is radiated. An antenna allows a signal to be amplified. This amplification is referred to as antenna *gain*, and it differs on factors like antenna type, elevation (ϕ) and azimuth (α). The *directivity* dictates the maximum value of the radiated energy.

The gain can be calculated from the *radiation pattern*. The radiation pattern is a function of the antenna type. For example, for an ideal dipole, the radiation pattern can be defined as (Carr, 2001):

$$E(\phi) = \cos(\phi) \tag{7.1}$$

Then, it follows that the gain in any direction is a multiplication of the directivity D and the radiation pattern (Carr, 2001):

$$G(\alpha, \phi) = D \cdot E(\phi) = D\cos(\phi) \tag{7.2}$$

Note, that the gain calculated in equation 7.2 depends only on the elevation ϕ. So, it is said the the gain for this antenna is isotropic in the azimuthal plane. In other words, energy is radiated equally for all azimuths.

Another interesting figure of merit is the *half-power beamwidth*, abbreviated HPBW . This parameter indicates at what angle the radiated energy is half the strength of the maximum radiated energy. This is an important parameter for this research, as it allows to determine in what directions the energy of an antenna is radiated. One wants to have the energy radiated in as many directions as possible, as the signals then ultimately cover a larger area. In order to find the HPBW , first consider that this is the bandwidth for which holds that the power equals half the maximum radiated power. The power is proportional to E^2. Hence:

$$E \quad = \quad \sqrt{P_{half}} = \sqrt{0.5} \tag{7.3}$$

The HPBW can then be found by inserting equation 7.3 into equation 7.1 and solving for ϕ. For the example antenna given in equation 7.1, this yields a HPBW of $45°$.

7.1.2 Selection

An ideal short dipole antenna is selected for signal transmission at the beacons. Because an ideal transmitter is assumed, inefficiencies can be ignored. Short dipoles are a subclass of the dipole antenna, for which it holds that $l << \lambda_s$. This implies that the length of the antenna, l, is significantly shorter than the wavelength of the signal to be transmitted (λ_s). The use of a normal dipole antenna is disadvantageous, because the length of a dipole antenna is typically equal to the wavelength or half of the wavelength of the signal. Consider that the signals, relevant in this study, have frequencies in the HF range, so $1MHz$ to $10MHz$. As such, the wavelength varies from $30m$ to $300m$. It is considered impractical to place beacons on Mars which has an antenna long enough to transmit HF wavelengths.

The short dipole antenna has a radiation pattern which is favorable for the use of the proposed beacon network. This antenna radiates isotropically in the azimuthal plane. This implies that energy is radiated with the same intensity for all azimuths (Balanis, 1996, p. 143). The radiation pattern in the elevation plane is described as follows (Balanis, 1996):

$$E(\phi) = \cos^2(\phi) \tag{7.4}$$

The gain then follows from rewriting equation 7.2 and combining with equation 7.4:

$$G = D\cos^2(\phi) \tag{7.5}$$

The directivity for a short dipole antenna is 1.5 (Carr, 2001). An inspection of equation 7.5 shows that the radiated energy is at a maximum at an elevation of $0°$. No energy is radiated at an elevation of $90°$. For example, at an elevation of $60°$, the gain is 0.375 or $-4.36dBi$.

The radiation pattern in the elevation plane is depicted in figure 7.1. From the figure it can be seen that the antenna radiates much of its power for a wide range of elevations. This is favorable, as the signals will be reflected by the ionosphere, and as such a large area will be covered by the signal from one beacon. Furthermore, little energy is radiated at high elevations. Again, this is favorable. Recall from section 4.2.2 that signals, transmitted with an elevation above the critical elevation will not be reflected by the ionosphere. Instead, they disappear into outer space. Hence, one wants to avoid wasting energy on signals transmitted in those directions. In conclusion, the short dipole is considered to be a suiting antenna for a first-order estimation of the performance of a ground-based beacon network.

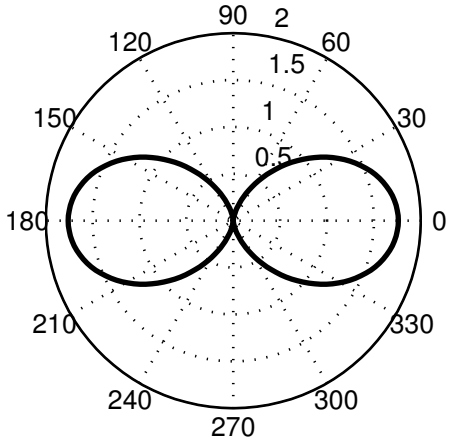


Figure 7.1: Radiation pattern for a short dipole antenna as a function of elevation. The radials indicate the gain in dB, compared to an isotropic antenna (Balanis, 1996).

7.2 Link budget

A link budget is calculated to determine the maximum acceptable path loss for a successful connection between a beacon and a mobile station. In order to establish how much path loss is acceptable, a couple of design questions must be answered:

- How much data must the link be able to transfer?
- What transfer reliability is required?
- What system hardware is available?

These questions need to be answered first, before a link budget can be established. As soon as the system hardware and the link requirements are determined, the calculation of the link budget follows naturally.

7.2.1 Communication basics

The amount of data, which can be transferred through a link, is determined by the *data rate* and the *channel bandwidth*. The data rate, abbreviated DR , states the number of bits which are transferred per second. The bandwidth, measured in Hz, is simply the difference between the upper and lower limit of the operating frequency in the link. There is a limit on the maximum error-free data rate, given a certain bandwidth. This limit, the *Shannon limit*, states that the maximum data rate which can be transmitted, given a certain bandwidth B, is (Larson and Wertz, 1999):

$$DR_{max} = B \log_2 (1 + SNR) \tag{7.6}$$

The SNR represents the *signal-to-noise* ratio. Evaluating equation 7.6 for a given data rate and bandwidth yields the lower bound on the SNR . When a SNR below this lower bound is used, no error free transmission is possible.

The data rate and bandwidth are also related through the *modulation method*. Modulation dictates how digital information is encoded into an analog signal. Several methods are available, like *frequency modulation*, *amplitude modulation* or *phase shift keying*. The latter encodes information by modifying the phase of an analog signal. A receiver then has to

measure the phase of the received signal, in order to extract the information of the signal (Larson and Wertz, 1999). Each modulation technique has its advantages and disadvantages, but such discussion is out of the scope of this work. Typically, the bandwidth is once or twice the data rate, depending on the modulation technique chosen (Zyren and Petrick, 1998, Table 1).

7.2.2 Link design

Because this work does not focus on mission hardware design, the required hardware is established based on some assumptions. First, it is assumed that the localization system is not needed for communication. Because of this assumption, only a low data rate is required. A suggestion is to have the beacons encode the local ionospheric conditions into the signal. This information can purposely be used by a mobile station to enhance a position estimate. As such, a small data rate would be required to provide for this option. Hence, it is assumed that a data rate of $5kbps$ is sufficient. Furthermore, a BPSK (*binary phase shift keying*) modulation technique is chosen. According to (Zyren and Petrick, 1998, Table 1), the typical bandwidth is twice the bit rate for this technique. Thus, the bandwidth for the selected frequency is $10kHz$.

Another design parameter that must be selected is the *bit error rate*, abbreviated BER . This parameter determines the transmission reliability of information through the link. The BER indicates the probability that a bit is lost during transmission (Larson and Wertz, 1999). So, for a BER of 10^{-3}, one in a thousand bits are lost during transmission. A BER of 10^{-7} is selected for the purpose of this work, as it a typical rate in many applications.

The last design parameter to select is the transmit power. The Mars Science Laboratory uses a UHF antenna with a transmit power of $9W$ (Makovsky et al., 2009). It seems therefore not unreasonable to select a comparable value. Hence, a transmit power of $10W$ is selected. It is assumed that this value is entirely within the technological limitations of present-day hardware, which is deployed on Mars. Furthermore it is assumed that a beacon can be designed, containing a power system adequate to provide a transmit power of $10W$ continuously.

Now, all the hardware and link design values have been selected. The next step is to calculate the gains and losses, by using these design values, so that ultimately the path loss margin can be retrieved.

7.2.3 Link calculation

The ultimate question of finding the path loss can be answered by calculating the link budget. The surplus in energy, which remains after adding all gains and subtracting all losses, is the amount of energy which can be used to overcome the path loss.

One primary element that has not yet been considered is the *noise power*, N. This measure indicates the lowest possible noise level for a certain communication system. The noise power depends on Boltzman's constant k, the system temperature T_{sys} and the bandwidth B, and is calculated as follows(Larson and Wertz, 1999):

$$N = kT_{sys}B \tag{7.7}$$

For a typical application, a system temperature of $290K$ can be assumed. Evaluating equation 7.7 in units of decibels then yields a noise power of $N = 163.98dB$ for a bandwidth of

$10kHz$. This is, however, the theoretical noise power for an ideal receiver. Typically, a receiver adds some noise due to imperfections in the hardware used. The receiver noise can be up to $15dB$ when cheap hardware is used. As such, it is assumed that the total receiver noise amounts to $10dB$.

Another factor, which should be taken into account, is the SNR . The SNR is a function of the received energy-per-bit to noise density, E_b/N_0, the bandwidth, and the data rate (Larson and Wertz, 1999; Zyren and Petrick, 1998):

$$SNR = (E_b/N_0)\frac{DR}{B} \tag{7.8}$$

The factor E_b/N_0 in itself is a function of modulation technique used and the BER . Without going too much into detail, it follows from Larson and Wertz (1999) that this factor is $E_b/N_0 = 11.5dB$ for a BER of 10^{-7}. The evaluation of equation 7.8 then yields a SNR of 7.06 dB. This value is cross-checked with equation 7.6 to see if the required data rate can be achieved. An evaluation of equation 7.6 shows that the maximum data rate is $DR_{max} = 30.1kbps$. The selected data rate is well within the limits.

Now, consider the values that are relevant in the link budget. For the transmitter, they are the transmit power, P_{TX} and transmitter gain G_{TX}. Then, for the receiver they are the receiver gain G_{RX}, receiver losses L_{RX} and receiver sensitivity P_{RX}. It is assumed that the receiver has a passive antenna. Lastly, the noise floor is indicated by NF, and the fade margin by M. The general formula for the link budget, written as a function of the path loss L_P is then (Larson and Wertz, 1999):

$$L_P = P_{TX} + G_{TX} - L_{TX} + P_{RX} + G_{RX} - L_{RX} + N - SNR - M \tag{7.9}$$

Table 7.1: Link budget for a short dipole antenna

Parameter	Unit	Value	Totals
Transmit power	W	10.0	
Data rate, DR	bps	5000	
bandwidth, B	Hz	$1 \cdot 10^4$	
BER	$-$	10^{-7}	
Transmitter power, P_{TX}	dBw	10.0	
Transmitter gain, G_{TX}	dBi	1.76	
Transmitter losses, L_{TX}	dB	-1.00	
Total	dB		**10.76**
Receiver sensitivity, P_{RX}	dBw	10.0	
Receiver gain, G_{RX}	dBi	1.76	
Receiver losses, L_{RX}	dB	-10.00	
Total	dB		**1.76**
Noise power, N	dB	163.98	
SNR	dB	-7.06	
Fade margin, M	dB	-30.0	
Total	dB		**126.92**
Path loss margin @ $\phi = 0°$	dB		**139.44**
Path loss margin @ $\phi = 75°$ (hypothetical)	dB		**119.64**

Note that the units in equation 7.9 are all in decibels, as this allows for an easy budget-like addition and subtraction of values. Converting the transmit power to decibels yields a transmit power of $10 \cdot \log_{10}(P_{TX}) = 10dBw$.

The budget is shown in table 7.1. It is found by evaluating equation 7.9, that a maximum path loss of 139.44 is acceptable. Note that this is the path loss for an elevation of $0°$. The path loss decreases with an increasing elevation angle, according to equation 7.5. Hence, the elevation angle at both the beacon and the mobile station should be taken into account in order to determine whether a link between a beacon and a mobile station can be established. To give an indication: in a hypothetical case where the elevation is $75°$ at both the beacon and the mobile station, the respective gains reduce to $G_{TX} = G_{RX} = -9.9dBi$. As a result, the path loss margin decreases severely. This effect is to be taken into account when determining whether a specific signal can be successfully received at a mobile station.

The discussion so far has focused on the link budget performance with the use of a short dipole antenna. The question is, if the maximum path loss can be increased. It is expected that the selection of an alternative antenna can yield better performance by providing a higher gain. This increases the maximum path loss. Ultimately, the coverage of a beacon is increased, as a higher path loss can be tolerated. As a result, a mobile station can successfully receive a signal from a specific beacon at a higher number of locations, hence increasing the coverage. Therefore, it is suggested for future work to research the optimal antenna for a ground-based beacon system.

7.3 Summary

It has been shown that a maximum path loss of $139.44dB$ can be tolerated when beacons are equipped with a short dipole antenna. This path loss is maximum at an elevation angle of $0°$ and decreases with increasing elevation. At an elevation of $75°$, the maximum path loss reduces to $119.64dB$. A suggestion for future research is to optimize the antenna design, such, that a higher maximum path loss can be provided. This ultimately leads to a better position estimate, as the coverage of beacons increases.

8 — Localization Technologies

Several technologies are available for localization based on a received radio signal, each exploiting a different parameter of the incoming radio signal. The most common parameters studied for localization are the amplitude, angle and timing of the signal as it reaches some receiver. However, each technology bears certain advantages and disadvantages. Hence, the question is, which technology yields the most accurate position estimate.

This chapter strives to give a qualitative discussion on the advantages and disadvantages of each technology. First, however, in section 8.1 an introduction in the principles of localization is given. Subsequently the various technologies relevant for this research are described. The following four technologies are studied:

- RSS : Received signal strength
- TOA : Time of arrival
- TDOA : Time difference of arrival
- AoA : Angle of arrival

These technologies are described in section 8.2. As the ionosphere is used for radio wave propagation, the beacon system inherently operates in a non-line-of-sight (NLOS) environment. Therefore each section concludes with a discussion on the NLOS effects on the usage of its respective technology. Finally a notion is given in section 8.3 on the GDOP , a metric used to identify the quality of the geometrical arrangement of the beacons, irrespective of the technology that they use.

8.1 Principles of localization

Localization in a ground-based beacon network is done by estimating geometric parameters between beacons (BC) and a mobile station (MS). Geometric parameters of interest can be distance or angle, which the mobile station derives from the received radio signal by reading out the signal parameters. An analytical model is applied to relate the distance between a beacon and a mobile station to these signal parameters. The measurements of radio signals from several beacons must be combined so that triangulation can be performed, leading to a position estimate. The dimensionality of the position determines the minimum number of beacons needed. At least three beacons are required to provide a solution in 2D-space, and four to determine a solution in 3D-space. The solution itself spans an area wherein the mobile station is positioned. The geometric parameter measured is either the distance from beacon n, $\hat{d}_n$, or the angle to a beacon, $\hat{\alpha}_n$. When the position is determined using distance measurements, the possible positions of a mobile station form a circle around a beacon. Hence, if three beacons are available, a unique position will exist in 2D-space where the three circles intersect. This position can be estimated by the product of measured distances of the individual beacons. The solution for such a position estimate is presented in figure 8.1.

The distances between a beacon and a mobile station in a ground-based beacon network are in the order of hundreds of kilometers, and therefore it cannot be assumed that the beacons and the mobile station lie on a flat plane. Thus, the positions are expressed in spherical

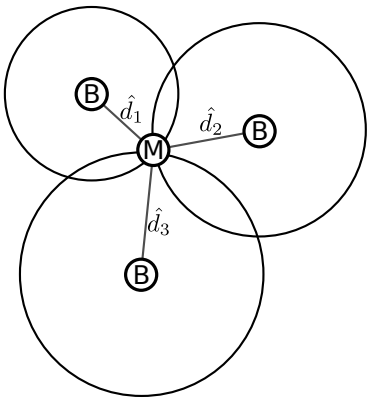


Figure 8.1: An illustration showing how a mobile station receives a 2D position estimate from three beacons. Estimated distances between beacons and a mobile station are indicated by $\hat{d}_n$. In the figure, the beacons are labelled B, and the mobile station M.

coordinates to account for the curvature of the planet. The distance between a mobile station and a beacon is expressed using the latitudes, β and longitudes, λ, of the beacon and mobile station. Refer to figure 5.4 for a definition of the coordinate system, in which these coordinates are defined. The separation between those two is expressed as a central angle $\delta\phi$, derived from the spherical law of cosines (Woan, 2000, eq.2.257):

$$\delta\phi = \cos^{-1}\left(\sin(\beta_{BC})\sin(\beta_{MS}) + \cos(\beta_{BC})\cos(\beta_{MS})\cos\left(|\lambda_{BC} - \lambda_{MS}|\right)\right) \quad (8.1)$$

$$d = \delta\phi R_M \quad (8.2)$$

Combining the angular separation from equation 8.1 with the planetary radius of Mars yields the *orthodromic* distance between any two points on a sphere, as expressed in equation 8.2. In this discussion, it is assumed that Mars is a perfectly spherical planet. This simplification can be made as the flattening of Mars is 0.589% of the major planetary radius[1]. Although Mars is more oblate than Earth[1], the flattening is still insignificant enough to be ignored for a first-order calculation of the performance metric. It is assumed that the inclusion of this parameter will not impact the comparison of the performance metrics for various environments and technologies, and therefore it is ignored.

Furthermore, the Northern hemisphere of the red planet is relatively smooth, containing mostly only slight terrain variations (Barlow, 2008). Thus, an assumption of a spherical planet is realistic for these smooth regions. However, the southern hemisphere is heavily cratered and therefore the terrain is rougher. Another example is the Tharsis region, a mountainous area. For these regions, one can not simply assume a flat surface. Radio signals in these regions will propagate longer or shorter, depending on the shape of the terrain. However, the impact of fluctuations in the terrain on localization estimation is a complex task in itself, and thus left for future research.

8.2 Technologies

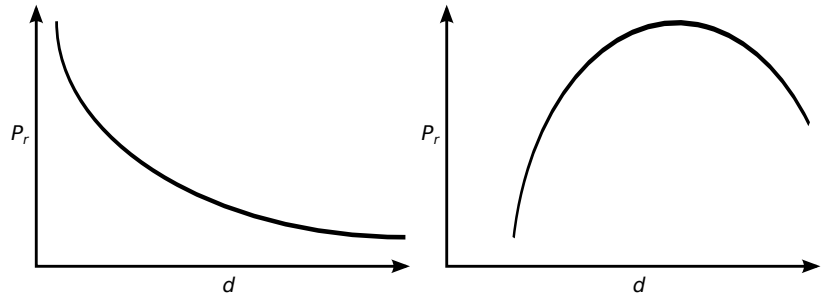
Each localization technology has certain advantages and disadvantages. This section treats every technology separately, demonstrating their main benefits and obstacles. As localiza-

[1] `http://nssdc.gsfc.nasa.gov/planetary/factsheet/marsfact.html`

tion is inherently done in an NLOS environment, the discussion focuses on the drawbacks caused by NLOS localization.

8.2.1 Received Signal Strength

Localization with received signal strength (RSS) is based on the dependence of the amplitude on the orthodromic distance between a mobile station and a beacon. Assuming this relationship is known, the distance between the two can be estimated by studying the amplitude, or signal strength, which the mobile station measures.



(a) RSS in a LOS environment *(b) RSS in a NLOS environment*

Figure 8.2: RSS as a function of distance. In a LOS environment, the received power P_r decreases logarithmically as the distance increases. In a NLOS ionospheric environment however, the relationship between received power and distance is parabolic (Kayton and Fried, 1997, fig.4.3).

The amplitude-distance relationship differs depending on the environment. Two common cases are depicted in figure 8.2. Interestingly, the received signal strength in a NLOS ionospheric environment is void at first, after which it increases. Only for increasingly large distances the signal strength decreases again. This behavior is caused by two effects: the skip distance and the ionospheric absorption. Consider that the distance a radio signal travels through the ionosphere depends on its initial launch elevation, as was shown in section 6.3. Radio signals with a very high launch elevation disappear into space as the Martian ionosphere cannot reflect them. A high launch elevation corresponds with a short distance between beacon and mobile station, and thus radio signals coming from a close by beacon cannot be received by a mobile station. This effect, the skip distance, explains the void in signal strength that appears in figure 8.2b.

Radio signals are absorbed and refracted by the ionosphere. Signals with a higher initial launch elevation penetrate more deeply into the ionosphere before they are reflected and therefore experience more severe absorption. As was previously mentioned, the orthodromic distance is inversely proportional to the initial launch elevation. At short distances, the signal strength is severely decreased. A decrease in launch elevation will increase the orthodromic distance and decrease the encountered absorption as the signal is reflected at a lower altitude. Thus, the signal strength increases with distance. Only at very large orthodromic distances, the path length starts playing a more significant role, resulting in decrease in received signal strength. This behavior is evident from figure 8.2b as well.

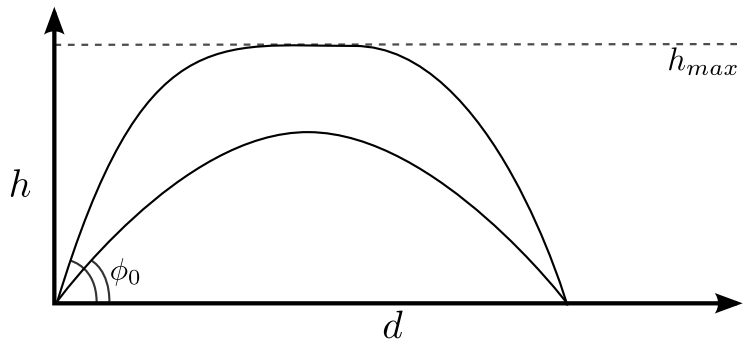


Figure 8.3: A schematic representation of multipath effects in an ionospheric NLOS environment. Rays with different launch elevation ϕ_0 are refracted differently. The altitude of the layer maximum is indicated by h_{max}. Some cases exist for which the orthodromic distance d of two rays with different launch elevation is identical.

8.2.2 Time of Arrival

Time-of-Arrival (TOA) technologies rely on the one-way time delay a signal experiences while traveling from beacon to mobile station. The delay time τ is an obvious product from the path length p and the speed of light c, where $\tau = p/c$. In NLOS environments like the ionosphere, it is hard to determine the path length, as the signals travel an indirect path via the ionosphere to the mobile station. Therefore, a linear relationship between the orthodromic distance, speed of light and time delay does not exist.

The time delay itself is derived from the difference of time at which the signal is broadcasted and at which it is received. This requires that the clocks in both the beacon and the mobile station are synchronized. This is not normally the case. Various effects can cause clocks to drift and desynchronize (Gentile et al., 2013). The effects of time synchronization are out of the scope of this research and therefore, it is assumed that all clocks are synchronized.

Multipath effects are profound in an ionospheric NLOS environment. Rays traveling from the same origin to the same destination may arrive at different time intervals. Such a case occurs when the rays are broadcasted with distinctly different launch elevations such that they are independently refracted by the ionosphere. They travel different paths but still arrive at the same location. In the situation, depicted in figure 8.3, two rays cover the same orthodromic distance but with different launch elevation and path length. The ray which is refracted close to the layer maximum is called the *Pedersen* ray, or high ray. When rays are refracted this close to the maximum, they travel long distances parallel to the maximum, thus greatly increasing the orthodromic distance. The normal ray has a lower launch elevation and is thus refracted differently. Ultimately, both rays reach the surface at the same distance, causing multipath effects. The Pedersen ray experiences significantly more absorption than the normal ray. Hence, it will show up as a secondary peak with lower signal strength on a time-versus- RSS diagram (Davies, 1969).

The severity of the multipath effects depends on the difference in time delay between individual rays. If the time difference is smaller than the time resolution of the system, the rays cannot be uniquely identified, and their amplitudes will be superimposed on each other (Gentile et al., 2013). However, the path differences in ionospheric propagation are in the order of tens of kilometers, causing a delay which is in the order of microseconds. These

delays are significant enough for modern electronics to detect and filter them.

8.2.3 Time Difference of Arrival

Localization using time difference of arrival (TDOA) compares the differences in delay times from signals received from multiple beacons. The difference in the delay is used to determine the orthodromic distance between the mobile station and a set of beacons. A position can be estimated by obtaining multiple measurements from different beacons. Each of these measurements yields a round-trip delay which corresponds to a solution for the position on a hyperboloid (Gentile et al., 2013). Combining the hyperboloids leads to a unique position estimation, assuming that there are sufficient measurements available. A position estimation in a 2D environment using three beacons is depicted in figure 8.4.

Two devices are required to establish a communication link which can be used for TDOA localization: an *initiator* and a *responder*. The mobile station performs the role of initiator by transmitting a signal requesting for position information. This signal is picked up by a beacon which responds with a second signal after a short processing delay. If more beacons are available, they will all respond at different time delays depending on their distance with respect to the mobile station.

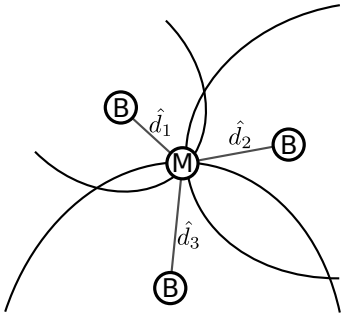


Figure 8.4: An illustration of a position estimate using a TDOA approach. The solutions of a position estimate are on a hyperboloid (Gentile et al., 2013).

The advantage of a TDOA over a TOA approach is that no synchronization of clocks is required. The total round-trip delay is directly measured at the mobile station. However, there is a delay involved when returning the transmitted signal. The electronics of the beacon acting as a responder require some processing time, naturally causing a response delay. The magnitude of this delay is required in order to determine the propagation delay. Therefore, information about the processing delay should be ultimately known to the initiator, a feat which can be achieved by encoding this information into the response signal.

The TDOA technology experiences the same problem as the TOA technology. As the rays travel forth and back to the initiator, however, the path they travel and thus the delay time is essentially doubled, disregarding the delay caused by the processing time. This is due to the laws of optics imposing reciprocity of the ray (see section 5.1.1). Thus, both the ray path and time delay will be equal in either direction.

8.2.4 Angle of Arrival

Angle-of-arrival (AoA) localization is performed by having the mobile station detect the elevation and azimuth angles of incoming rays. Three angles are required to uniquely es-

timate the position of a mobile station on a spherical surface. Figure 8.5 illustrates how a position is estimated on a flat world using two elevation angles. The same approach is used for a position estimate on a sphere, like Mars, but with the addition of a third beacon. The angles can be detected by making use of a multi-antenna array such as a phased array antenna. The precision by which the angles are measured depends on the number of elements in the array. Thus, it is evident that an array with a large number of elements is required for precise angle detection.

The elevation and azimuth of the incoming ray are estimated by measuring the phase differences between individual elements of the array. The ray geometry is such that, assuming a flat planet, the elevation angle received at the mobile station is equal to the launch elevation (Kelso, 1964, p.206). This assumption holds when extended to a spherical planet, when the spherical angle between a beacon and a mobile station is included. As a result, the only unknown parameter is the azimuth angle. The assumption does not hold when the altitude of the mobile station and a beacon differs significantly, such as in mountainous terrain. However, the northern hemisphere of Mars is known to be relatively flat such that a spherical planet can be assumed in this case. Ignoring the altitude of a mobile station, the measurement of elevation angles from three beacons is sufficient to determine the position of that mobile station on a spherical planet. Hence, in this work, the AoA approach measures the elevation angles of incoming signals.

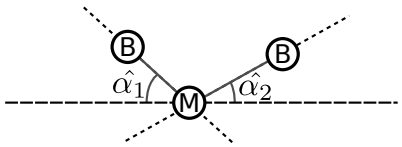


Figure 8.5: An illustration of AoA localization using two beacons, located at estimated azimuths $\hat{\alpha_1}$ and $\hat{\alpha_2}$ w.r.t. the mobile station respectively. Solutions for the position of the mobile station w.r.t each beacon lie on a line.

An advantage over all other methods is that AoA systems do not require any form of clock synchronization at all, as the estimates are purely based on incoming angle. The main disadvantage of AoA -based technologies is that multipath effects heavily degrade the performance of these systems. In mountainous terrain it is not inherently clear from which direction the original ray is coming, as rays can be reflected on oblique surfaces. Therefore, it is advisable to combine AoA technologies with others such as RSS or TOA .

8.3 Dilution of precision

The *geometric dilution of precision* (GDOP) is a metric which rates the quality of the geometric arrangement of the beacons in a positioning system, and is subsequently used as a performance metric for a position estimate. The GDOP combines the ranging errors in both position (PDOP) and time (TDOP). The PDOP measure relates the arrangement of the beacons in a system to the accuracy of the position estimate in each dimension (Kayton and Fried, 1997). The best results are obtained if the beacons are separated with a wide angular separation, as seen from the mobile station. When the beacons are placed roughly in the same direction, as seen from the mobile station, the accuracy of the position estimate is degraded. This effect is depicted in figure 8.6.

The PDOP metric is derived from the variances of the position estimate in each dimension. Assuming the variance of the distance measurement itself, σ_d^2, is equal for each measurement, the PDOP is defined by Kayton and Fried (1997) as:

$$PDOP = \sqrt{\frac{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}{\sigma_d^2}} \tag{8.3}$$

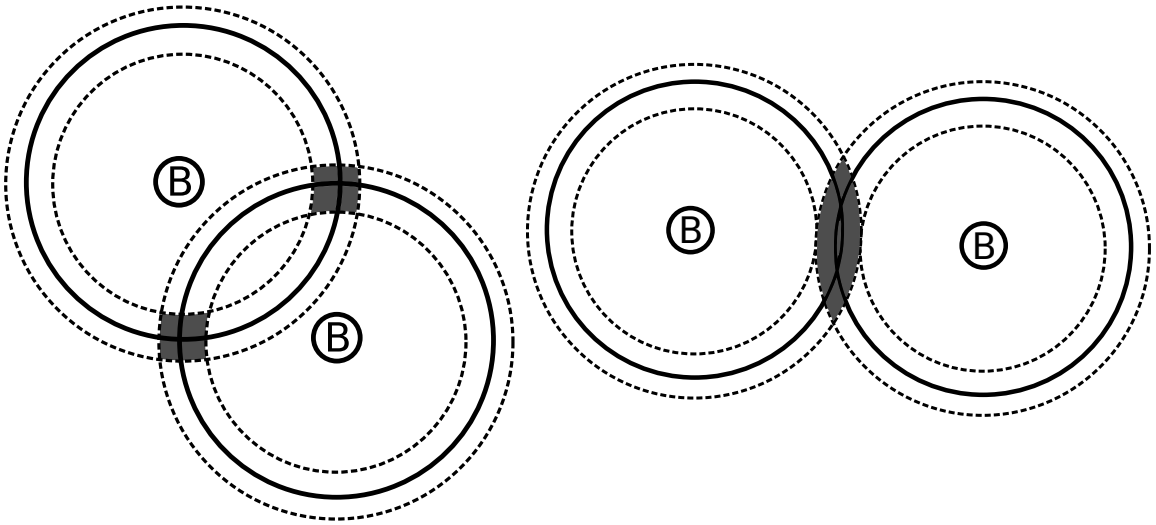
In a 2D-space, the metric can also be expressed by ignoring the variance in the z-direction. This yields the HDOP, the *horizontal dilution of precision*:

$$HDOP = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{\sigma_d^2}} \tag{8.4}$$

The measure for the GDOP itself is found by combining the PDOP and TDOP:

$$GDOP = PDOP + TDOP \tag{8.5}$$

An estimate is considered of high quality when the value generated from equation 8.5 is around unity. Estimates with GDOP < 1 are considered to be excellent, while values in the range of $1 - 3$ are considered good. Values from $3 - 6$ yield an acceptable position estimate, and any estimate with a GDOP higher than 6 is considered to be unacceptable (Bossler and Campbell, 2010; Kayton and Fried, 1997). Ultimately, the lower the GDOP, the better.



(a) An arrangement which yields a precise position estimate as the solution area is small. This is therefore an efficient arrangement of beacons, although not unique.

(b) An arrangement which yields an inaccurate position estimate as the solution area in one direction is large.

Figure 8.6: The GDOP is illustrated with two different arrangements. Solid circles indicate the solutions of the position estimate of a mobile station with respect to one beacon, while error bounds are indicated by dotted circles. Possible positions of the MS are depicted by a gray area. Note that in figure 8.6a, two areas exist. Hence, a third beacon is required to estimate a unique position.

In the scope of this research the GDOP metric is therefore used to indicate the quality of the obtained position estimates. By using the value of the GDOP found for a specific position estimate, one can determine if that estimate should be discarded, if the GDOP is too low. When this happens, no position estimate is available at this position. Hence there is no *coverage* at this position. This approach ultimately allows to determine at which positions a mobile station has coverage, and where not.

8.4 Summary

Each of the four technologies, treated in this chapter, brings different advantages and disadvantages. First, the RSS approach is subject to absorption in an NLOS environment. It was shown that the absorption increases with distance, and maxes out at some distance away from a beacon. The TOA approach suffers from synchronization issues between individual beacons. This issue is solved by the TDOA approach. In this approach, a mobile station sends a signal first to all beacons, upon which they respond. The AoA approach can yield a position estimate with one beacon less than the other approaches. However, mountainous terrain can degrade the position estimate. All approaches are effected by the NLOS environment of the ionosphere.

The GDOP is a powerful metric to determine the quality of a position estimate. This metric can be used to determine when a position estimate should be discarded. Ultimately, the GDOP can be used to determine the coverage. A localization system provides no coverage at a certain position when the estimate at this position is discarded.

9 — Estimation Model

The precision value can be used to determine and compare the performance of the localization technologies, hence proving to be an extremely useful metric. It is this metric that gives insight in the precision that a localization system can provide under given circumstances. Ultimately, the establishment of a quantitative performance metric is the basis for a feasibility study and allows for the hypothesis to be asserted. The performance metric follows from the evaluation of the *Cramér-Rao Lower Bound* (Kay, 1993), abbreviated CRLB , a powerful measure to indicate the minimum variance of any unbiased estimator. The derivation of this measure eventually leads to a numerical value of the performance metric, given a certain ionospheric environment.

This work aims to determine the *precision* which a positioning system can provide. The precision indicates the spread of a position estimate at one standard deviation. On the other hand, the *accuracy* dictates the deviation from the mean. The lower bound provided by the CRLB is infinitely accurate by its very nature, as it is the lower bound for an unbiased estimate. Thus, the goal is to express the performance of a positioning system in terms of precision, measured in meters.

This chapter focuses on deriving the expression for the performance metric for various positioning systems. Each of these positioning systems uses one specific localization technology. First, an overview is given of the assumptions that have been made so far. This is done in section 9.1. Next, a method is described to filter incoming signals so that accurate measurements can be obtained, in section 9.2. Finally, section 9.3 treats a general approach to the derivation of the performance metric. In this section, the general form of the CRLB for any localization technique is derived. Ultimately, this general form leads to the *characteristic error*. The characteristic error is the parameter which quantifies the performance of a position estimate. In this case, that is the precision of a position estimate.

9.1 Basic assumptions

Several assumptions have been made to achieve the results presented in this and subsequent chapters. This section gives a recap of the assumptions made so far, plus an argumentation about their validity.

The ionosphere is modeled as a Chapman-profile. The electron number density in the ionosphere follows a Chapman profile, a simplification which in itself makes a couple of assumptions. First, the atmosphere is assumed to consist of a single gas. Second, the neutral density decays exponentially with increasing altitude. The Chapman profile is described in more detail in section 4.1.2.

Ionospheric fluctuations are negligible. The Chapman-profile defines an electron density profile which is purely dependent on altitude and SZA . Temporal fluctuations are ignored, because their timescales are several orders of magnitude larger than the timescales relevant for radio waves. This assumption, too, is elaborated in section 4.1.2.

Mars does not have a strong magnetic field. In this work, it is assumed that the influence of Mars' magnetic field is insignificant. For further reference, see section 4.2.1.

Mars is a simplified to a sphere. It is assumed that Mars is a perfectly spherical planet, in order to simplify calculations with respect to the performance metric. The oblateness of the planet is relatively small. Furthermore, the Northern hemisphere consists mainly of smooth terrain. Section 8.1 treats this assumption in more detail.

Effects of beacon location can be simplified. For beacons placed close to the Subsolar point, the Solar zenith angle has a negligible impact on the calculations. At the Subsolar point, the Solar zenith angle is null by definition. For small SZA angles, a simplification can be made by using a small-angle-approximation, which states that $\cos(\alpha) \approx 1$ for $\alpha < 5°$. Without loss of generality, this simplification can be applied on Chapman theory, which frequently uses $\sec(\chi)$ to indicate the effects of the Solar zenith angle on the electron number density (see section 4.1.2). Thus, the simplification $\sec(\alpha) \approx 1$ is used. Placing a beacon close to the Subsolar point makes that the radio signals, originating from that beacon, follow an equivalent behavior to radio signals originating from the Subsolar point itself. Therefore, it is assumed that the location of the beacon with respect to the Subsolar point is irrelevant, as long as the central angle between the two is less than $5°$. Beacons who are placed further away will need to take this effect into account, however.

9.2 Signal filtering

Radio signals are inherently diffuse when they are received, and signal filtering is required to filter out noise. The diffusion is caused by multipath effects in the ionosphere, who constitute to varying propagation times and distortion in other signal parameters. Recall from section 8.2.2 that rays are refracted differently depending on their initial launch elevation. As was seen in figure 8.3, a distinction can be made between normal rays and *Pedersen* rays, which take a longer path. A mobile station at any location on Mars will thus receive multiple signals from the same beacon at different delay times. This is depicted in figure 9.1, and it affects the way in which the mobile station interprets the combined signal. Consider that each ray will have a different propagation time, signal strength, and angle of arrival. If a mobile station would use all the signals, received over a long time, for position estimation, the result would be very imprecise as there is a lot of inconsistency in the received signal properties. Hence, a method of filtering is applied to enhance the position estimate by reducing the inconsistence caused by the various rays.

The proposition is that a position estimate can be extracted from only the first signal from each beacon. Hereby it is assumed that beacons broadcast signals in a discontinuous fashion, so that an initial signal exists in the first place. Furthermore, it is assumed that the signals which arrive first must have taken the shortest path, as the distance traveled is linearly related to the speed of light. Therefore, they are the signals which have been least affected by multipath effects. The goal is to filter out all the other signals at the mobile station, so that the filtered signal only contains the rays which took the shortest path to the mobile station. This is done by windowing the received signal in the time domain. Many windowing methods are available. In this work, a Dirichlet window (also called a rectangular window) is selected to filter out the unwanted rays. It is defined as follows (Woan, 2000,

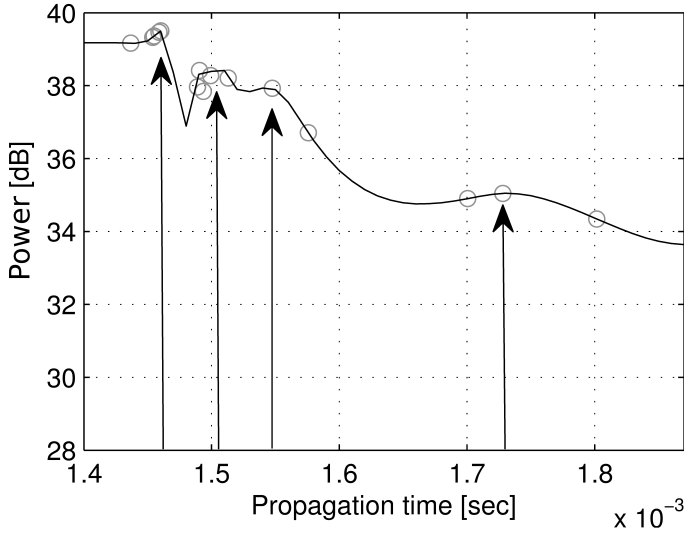


Figure 9.1: An example of an inverted power profile as received by a mobile station. Individual signals are indicated by circles. The clusters they form, indicated by arrows, are a result from the multipath effects. The first cluster which is received, relates to the reception of normal rays. Subsequent clusters relate to Pedersen rays.

p. 50):

$$w(\tau) = \begin{cases} \tau_0 < \tau < \tau_0 + \Delta t & 1 \\ \text{otherwise} & 0 \end{cases} \tag{9.1}$$

In equation 9.1, the propagation time of the initial signal is denoted by τ_0. The length of the observation interval is Δt. The window function is denoted by $w(\tau)$. Essentially, this window filters out all the rays which arrive outside the observation interval at the mobile station. Note that the Dirichlet window, although straightforward, has several downsides compared to other windowing methods (Harris, 1978). However, the application of a proper signal filter is outside of the scope of this work. Hence, an ideal Dirichlet window without side effects is assumed.

To summarize, an ideal Dirichlet window provides an easy method of countering the multipath effects which are inherent in an ionospheric environment. The results on position estimation which are presented in this work are all filtered using an ideal Dirichlet window.

9.3 Estimating a position using the CRLB

The precision of the position estimation of a mobile station depends on its ability to measure and interpret radio signals incoming from the beacons. Several aspects of the incoming signal can be exploited. However, the measurements unavoidably contain some noise, which causes that the position estimate is not infinitely precise. The precision of a position estimate, subject to noise, can be evaluated with the CRLB . The CRLB states that the variance of some unbiased estimator is at least as high as the inverse of the Fisher information (Kay, 1993; Qi and Kobayashi, 2002). In this case, the parameter to be estimated is the position of a mobile station, measured in latitude and longitude. The CRLB gives a lower bound on the

covariance of a position estimate which can be obtained. Thus, for an unbiased estimator, no lower covariance is possible than the covariance described by the CRLB .

In a spherical world, the parameters to be estimated are the latitudinal and longitudinal coordinates. Together, these make up the *parameter vector*. Hence the parameter vector is defined as $\boldsymbol{\theta} = [\beta, \lambda]^T$. The true position of the mobile station, relative to a beacon, is defined as $\boldsymbol{\theta} = [\beta, \lambda]^T$. The estimated position is expressed as $\hat{\boldsymbol{\theta}} = [\hat{\beta}, \hat{\lambda}]^T$. A mobile station estimates its distance $\hat{d}$ to any one beacon by measuring some parameter of the signal. In this chapter, that measurement parameter is denoted as τ. The measurement parameter can come from any one of the technologies, described in the previous chapter. For example, in an TOA -approach, the time is measured. An RSS -approach measures the signal strength, whereas the AoA -approach measures the angle of arrival.

It is assumed that all measurements performed by the mobile station are independent and identically distributed. Furthermore, it is assumed that all noise is Gaussian distributed with zero mean. Then, the measured parameter is equal to a function of the parameter vector $f(\theta)$, embedded in zero-mean Gaussian noise $\boldsymbol{w}$:

$$\boldsymbol{\tau} = f(\boldsymbol{\theta}) + \boldsymbol{w} \tag{9.2}$$

The function of the parameter vector, $f(\boldsymbol{\theta})$, relates the estimated position of the mobile station to the measurement parameter. The output of this function is the time of arrival, $\boldsymbol{\tau}$. This can be clarified by an example. In a TOA -approach, the estimated time of arrival increases with distance from a beacon. This relationship can be modeled. The observed time of arrival, in this case $\boldsymbol{\tau}$, is the addition of the estimated time of arrival, $f(\boldsymbol{\theta})$, with some Gaussian noise. The likelihood that the parameter vector $\boldsymbol{\theta}$ takes some value, is described in the probability density function (PDF):

$$P(\boldsymbol{\tau}|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-1}{2\sigma^2}(\boldsymbol{\tau} - f(\theta))^2\right] \tag{9.3}$$

The subsequent approach, as stated by Kay (1993), shows how to obtain the CRLB from the measurement parameter. The CRLB itself equals the inverse of the expectation of the measurement parameter. The expectation of $\boldsymbol{\tau}$ can be found by derivation of the conditional PDF from equation 9.3 with respect to the measurement parameter:

$$E(\boldsymbol{\theta}) = E\left[\frac{\partial^2 lnP(\boldsymbol{\tau}|\boldsymbol{\theta})}{\partial^2 \boldsymbol{\theta}}\right] \tag{9.4}$$

The CRLB can be found from 9.4. The unbiased estimator in this case is $\hat{\boldsymbol{\theta}}$, so that:

$$var(\hat{\boldsymbol{\theta}}) \geq \frac{1}{-E\left[\frac{\partial^2 lnP(\boldsymbol{\tau}|\boldsymbol{\theta})}{\partial^2 \boldsymbol{\theta}}\right]} \tag{9.5}$$

Equation 9.5 shows the CRLB , which states that the minimum variance is either equal or higher than the inverse of the Fisher information. This equation may be rewritten in terms of the *Fisher information matrix*, which in itself is defined by:

$$\boldsymbol{I}(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 lnP(\boldsymbol{\tau}|\boldsymbol{\theta})}{\partial^2 \boldsymbol{\theta}}\right] \tag{9.6}$$

This turns the expression for the minimum variance of an unbiased estimator from equation 9.5 into:

$$var(\hat{\boldsymbol{\theta}}) \geq \frac{1}{\boldsymbol{I(\theta)}} \tag{9.7}$$

Concluding, equation 9.7 expresses the lower bound of the precision on the position estimate.

The measurements in the measurement vector $\boldsymbol{\tau}$ differ for each beacon whose signal is measured. The derivation of the CRLB can be extended to make use of multiple measurements. These measurements are defined for each beacon independently in a measurement vector $\boldsymbol{\tau}$, representing the propagation time:

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_n \end{bmatrix} \tag{9.8}$$

Next, the Fisher information matrix is required to ultimately compute the CRLB . In this case, it can be obtained by deriving the measurement vector from equation 9.8 with respect to each element in the parameter vector $\boldsymbol{\theta}$. Essentially, the measured propagation time for each beacon is derivated with respect to the latitude and longitude. This is achieved using a Jacobian J:

$$J = \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial \tau_{\mathbf{1}}}{\partial \beta} & \frac{\partial \tau_{\mathbf{1}}}{\partial \lambda} \\ \frac{\partial \tau_{\mathbf{2}}}{\partial \beta} & \frac{\partial \tau_{\mathbf{2}}}{\partial \lambda} \\ \vdots & \vdots \\ \frac{\partial \tau_{\mathbf{n}}}{\partial \beta} & \frac{\partial \tau_{\mathbf{n}}}{\partial \lambda} \end{bmatrix} \tag{9.9}$$

Then, the Fisher information matrix $\boldsymbol{I}(\boldsymbol{\theta})$ can be written as:

$$\boldsymbol{I}(\boldsymbol{\theta}) = \boldsymbol{J}\boldsymbol{M_P}\boldsymbol{J^T} \tag{9.10}$$

The measure of precision M_P in equation 9.10 is the inverse of the covariance matrix $\sum$, spanned for the number of beacons in the system:

$$\boldsymbol{M_P} = \sum^{-1} = diag(\sigma_1^{-2}, \sigma_2^{-2}, \ldots, \sigma_n^{-2}). \tag{9.11}$$

The use of the standard deviations in equation 9.11 requires some clarification. The standard deviations of the measurement parameters τ_n from the parameter vector $\boldsymbol{\tau}$ are indicated with σ_n. A mobile system measures some signal parameter τ, and relates it to the distance to a beacon with an analytical model fit, as will be described in detail in chapter 10. The measurement unavoidably contains noise, hence yielding a standard deviation in the analytical model. The standard deviations indicate dispersion caused by noise from the ionosphere. The dispersion in the model fit is the source of the standard deviations in equation 9.11. The simulation uses a Monte-Carlo approach to model the disturbances from the ionosphere, as described in section 5.3. The standard deviations in equation 9.11 are independent as they relate to different beacons. Another way to visualize this, is to consider figure 8.6. The distance between the dotted lines (bound of the uncertainty) and the solid lines (distance) indicates the standard deviation of the uncertainty.

The CRLB itself is defined as the inverse of the Fisher information matrix:

$$CRLB = \mathbf{I}(\boldsymbol{\theta})^{-1} \tag{9.12}$$

Combining equations 9.10 to 9.12 yields the CRLB expressed in terms of a set of nonlinear equations, which can be found by computing the Jacobian given in equation 9.9. The CRLB takes the form of a square matrix with its dimensions set to the number of parameters in this matrix.

$$CRLB = \left(\boldsymbol{J}\boldsymbol{M}_{\boldsymbol{P}}\boldsymbol{J}^{T}\right)^{-1} \tag{9.13}$$

In this case, the CRLB takes the shape of a $2x2$ matrix. The variances of the longitude and latitude can be found on the diagonals, so $CRLB_{1,1}$ and $CRLB_{2,2}$ respectively. In a geometrical sense, these elements define the semi-major and semi-minor axis of an error ellipsoid. The off-diagonal elements, $CRLB_{1,2}$ and $CRLB_{2,1}$ subsequently determine the rotation of the ellipse.

Ultimately, the parameter of interest for determining the performance of a positioning system is the precision, $\hat{p}$. This is the *characteristic error* which is derived from the diagonals of the CRLB matrix.

$$\hat{p} = \sqrt{var(\hat{\beta}) + var(\hat{\lambda})} \tag{9.14}$$

The precision given in equation 9.14 is a measure of performance for the system. The resulting quantity is measured in radians, as the measurements relate to angular separations between mobile station and beacon. In order to obtain a meaningful measure of precision, the characteristic error can be translated to units of meters by simply multiplying with Mars' equatorial radius R_M:

$$\hat{p} = R_M\sqrt{var(\hat{\beta}) + var(\hat{\lambda})}\ [m] \tag{9.15}$$

The result in equation 9.15 gives the precision, in meters, for one specific position of the mobile station with respect to a set of beacons. However, this position might change, yielding different values for the precision. It is therefore of interest to explore the lower bound on the precision for all possible positions of the mobile station with respect a set of beacons. This requires the computation of the aforementioned method for all possible positions, and will be left for the next chapter.

9.4 Summary

In this chapter, an estimation model is provided which determines the precision of a positioning system using the CRLB . The CRLB is the smallest covariance which is attainable by any uunbiased estimator that uses this model. Some assumptions and simplifications have been made. First and foremost, Mars is assumed to be a spherical planet, an assumption which aids and simplifies the calculation of the CRLB . Estimation models which include terrain variations have been left for further work. Furthermore, a simplified ideal filter is used to filter signals as they arrive at the mobile station. This method allows to increase the precision even further, by decreasing noise caused by multipath effects.

The computation of the CRLB aids in translating the distance measurements to a lower bound on the precision. Such can be achieved using equations 9.8 to 9.15. To explore the variations in precision for all positions of a mobile station, the position space must be explored, as will be done in the next chapter.

10 — Comparison of Technologies

The performance metric, found in the previous chapter, is a valuable indication to compare different localization technologies, as the performance of the localization system depends on the technology used. Comparing the performances of the previously selected localization technologies gives insight in which technology performs best under a given circumstance.

This chapter focuses on establishing a method of comparison of the performance metric for the nominal scenario. See chapter 6 for a definition of this scenario. The procedure described in chapter 9 is performed for each localization technology in order to obtain its respective performance under certain nominal conditions. First, an analytical expression, based on the ray-tracing simulation, is derived for the parameter which will be measured by a mobile station. The analytical expression relates the orthodromic distance to the parameter which is measured. It aids in computing the performance metric for various positions of the mobile station. This is done by scanning the entire space of possible positions of the mobile station with respect to a certain set of beacons. The performance metric is computed for each position, the result of which is shown in a heatmap. This heatmap demonstrates the precision of the position estimate which a mobile station can receive at any location.

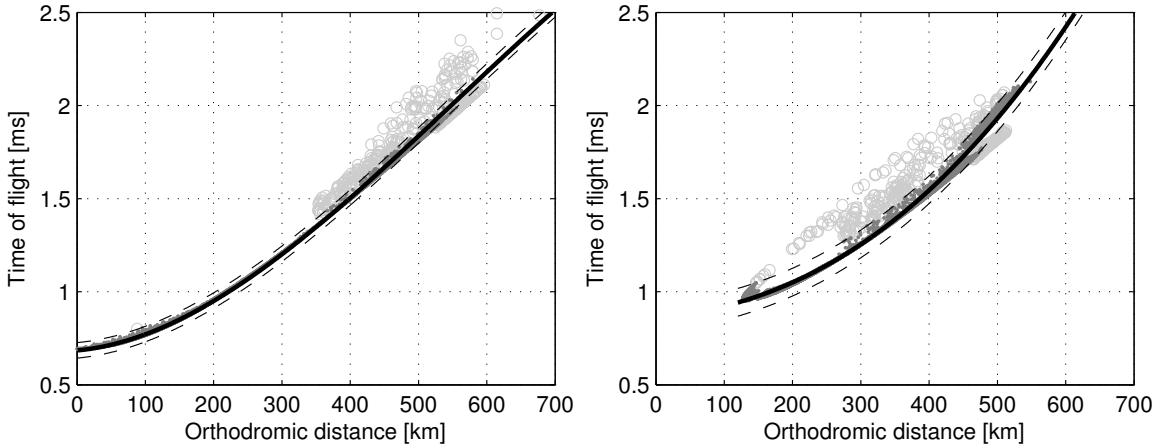
The performance is expressed as the *mean estimated precision*, denoted as $\mu_{\hat{p}}$ with a unit in meters. The mean precision is determined by taking the mean of all precisions, found in the heatmap, inside the area bounded by the beacons. The beacons are points of a square with sides of 10° in latitude and longitude. The space inside the square defines the area. Inside this area, the precision is measured for each position of the mobile station. Position estimates outside the square are ignored. The latter is done under the assumption that a mobile station always stays inside the area bounded by the beacons. The effect of this assumption is merely that a localization system must be set up such that the beacons surround the area for which one wants to provide positioning information. The performance of the area inside the square is averaged, yielding $\mu_{\hat{p}}$.

In the subsequent comparison, a beacon system consisting of four beacons is chosen. A minimum of three beacons is required to uniquely estimate a position, as was shown in chapter 8. A fourth beacon is added to ensure that there are always three beacons in range. A subsequent chapter will deal with beacons which are out of range or otherwise unavailable. The only limitation, which is taken into account for this comparison, is the skip distance. For frequencies above the plasma frequency, signals of nearby beacons cannot be received if the orthodromic distance to them is lower than the skip distance. Section 6.3.2 explains this effect in more detail.

The performance is evaluated separately in sections 10.1 to 10.3 for the TOA, RSS and AoA approaches, respectively. A discussion and comparison is presented in section 10.4.

10.1 Time of Arrival

The system performance in a system using TOA is determined by finding the CRLB of the precision using measurements obtained from data of the ray-tracing simulation. In a TOA



(a) $f = 2.5MHz, \bar{\sigma} = 0.21ms$ *(b)* $f = 4.5MHz, \bar{\sigma} = 0.38ms$

Figure 10.1: A least-squares polynomial fit of the propagation time as a function of the orthodromic distance for various frequencies. Solid black lines indicate the least-squares fit, whereas dashed black lines indicate the 95% *confidence level of the fit. Light gray circles indicate ignored data. The fit is based on the dark gray points. The outlier rejection is done using the window fit described in section 9.2. The mean standard deviation,* $\bar{\sigma}$*, increases with frequency*[1].

approach, the distance between a beacon and a mobile station follows from the time the rays take to propagate from one to the other. As is described in chapter 8, localization is done in a NLOS environment and as a result the distance is not simply dependent on the propagation time and speed of light.

To find the CRLB for a TOA system, one wants to measure the propagation time τ and relate it to the distance between beacon and mobile station. Therefore, in this case, propagation time is the measurement parameter. The relationship can be found using simulation data, as depicted in figure 10.1. In this figure, grey data points indicate at what delay and distance from the origin a ray collided with the surface. At this point, it is imagined that the ray is received by a mobile station. Therefore, the data points indicate propagation times of rays received by a mobile station at a certain distance away from a beacon. Herein, a distinction has been made between data which is discarded (light gray in the figure), and data which has been preserved for model fitting (dark gray). The distinction is performed using the approach previously described in section 9.2. An analytical expression for the relationship can be found by fitting the data to a polynomial using a least-squares fit. It is found that a cubic polynomial closely represents the relationship between the orthodromic distance d_n, defined in equation 8.2, and propagation time τ:

$$\tau = ad_n^3 + bd_n^2 + cd_n + g \ [ms] \tag{10.1}$$

In equation 10.1, the constants a, b, c and g vary for the frequency used and the composition of the environment. Therefore, equation 10.1 needs to be re-evaluated for different frequencies and scenarios, leading to different performance metrics. Figure 10.1 demonstrates the least-squares fit obtained for ray frequencies of $2.5MHz$ and $4.5MHz$. Table 10.1 gives an overview of the constants obtained to create the fits for several frequencies.

Furthermore, the standard deviation of the least-squares fit, which varies with frequency too, is derived from the confidence level of the fit. The fits made in figure 10.1 use a con-

[1]The standard deviation is different for each beacon. Nevertheless, it is found that this difference is less than 1%. Hence, the *mean* standard deviation is used as a simplification.

Table 10.1: Model fit parameters for various frequencies where TOA is used. The polynomial fit has the shape $\tau = ad_n^3 + bd_n^2 + cd_n + g$. The data is obtained using the nominal scenario.

f	$2.5MHz$	$3.5MHz$	$4.5MHz$	5.5 MHz
a $[ms\,km^{-3}]$	$-3.405 \cdot 10^{-12}$	$-4.339 \cdot 10^{-12}$	$1.55 \cdot 10^{-12}$	$-2.299 \cdot 10^{-11}$
b $[ms\,km^{-2}]$	$5.626 \cdot 10^{-9}$	$6.424 \cdot 10^{-9}$	$2.92 \cdot 10^{-9}$	$2.83 \cdot 10^{-8}$
c $[ms\,km^{-1}]$	$3.482 \cdot 10^{-7}$	$4.28 \cdot 10^{-8}$	$-2.97 \cdot 10^{-7}$	$-8.71 \cdot 10^{-6}$
g $[ms]$	$6.85 \cdot 10^{-4}$	$7.60 \cdot 10^{-4}$	$8.63 \cdot 10^{-4}$	$2.00 \cdot 10^{-3}$
1σ	$2.12 \cdot 10^{-2} ms$	$2.43 \cdot 10^{-2} ms$	$3.80 \cdot 10^{-2} ms$	$1.73 \cdot 10^{-2} ms$
Skip distance	$0km$	$0km$	$127.1km$	$271.9km$

fidence level of 95%, indicated by dashed lines. The standard deviation for any given frequency can subsequently be derived from these confidence levels.

A lack of data points for lower distances in figure 10.1 demonstrates that the minimum distance for which measurements are available increases with frequency. This is the effect of the skip distance, which was previously discussed in section 6.3.2. The skip distance for $f = 2.5MHz$, shown in figure 10.1a, is non-existent, whereas the skip distance for $f = 4.5MHz$ (figure 10.1b) equals $127.1km$. These results agree with the results found for the MUF in section 6.3.2.

The distance is represented as an orthodromic distance, which has been defined before in equation 8.2 and can be combined with equation 10.1 to yield the full expression for the measurement vector. The analytical expression in equation 10.1 only describes the propagation time between the mobile station and one specific beacon. In order to get a full expression of the measurement vector, all beacons in the system need to be taken into account. Assuming n beacons available, this leads to the following expression:

$$\tau = \begin{bmatrix} ad_1^3 + bd_1^2 + cd_1 + g \\ ad_2^3 + bd_2^2 + cd_2 + g \\ \vdots \\ ad_n^3 + bd_n^2 + cd_n + g \end{bmatrix} \tag{10.2}$$

The next step leading to the CRLB for a TOA approach is to derive the Jacobian of the measurement vector, previously defined in equation 9.9. It is not of interest to describe the complete computation of all the partial derivatives here. A general derivation for both $\frac{\delta\tau}{\delta\beta}$ and $\frac{\partial\tau}{\partial\lambda}$ can be found in appendix A.1. The resulting partial derivative with respect to the latitude for a certain beacon n is found in equation A.8:

$$\begin{aligned} \frac{\partial\tau}{\partial\beta} &= \left(3ad_n^2 + 2bd_n + c\right) \frac{R_M}{\sqrt{1-x^2}} \\ &\quad \cdot \left(\cos(\beta_{BC})\sin(\beta_{MS})\cos(\Delta\lambda) + \sin(\beta_{BC})\cos(\beta_{MS})\right) \end{aligned} \tag{10.3}$$

Similarly, for the longitude, found in equation A.13:

$$\frac{\partial\tau}{\partial\lambda} = \left(3ad_n^2 + 2bd_n + c\right) \frac{R_M}{\sqrt{1-x^2}} \cos(\beta_{BC})\cos(\beta_{MS})\sin(\Delta\lambda) \tag{10.4}$$

The measure of precision, defined in equation 9.11, consists of the standard deviations of the analytical fit. These are given in table 10.1. It is found that the standard deviations are equal within 1%. The measure of precision for various frequencies subsequently yields:

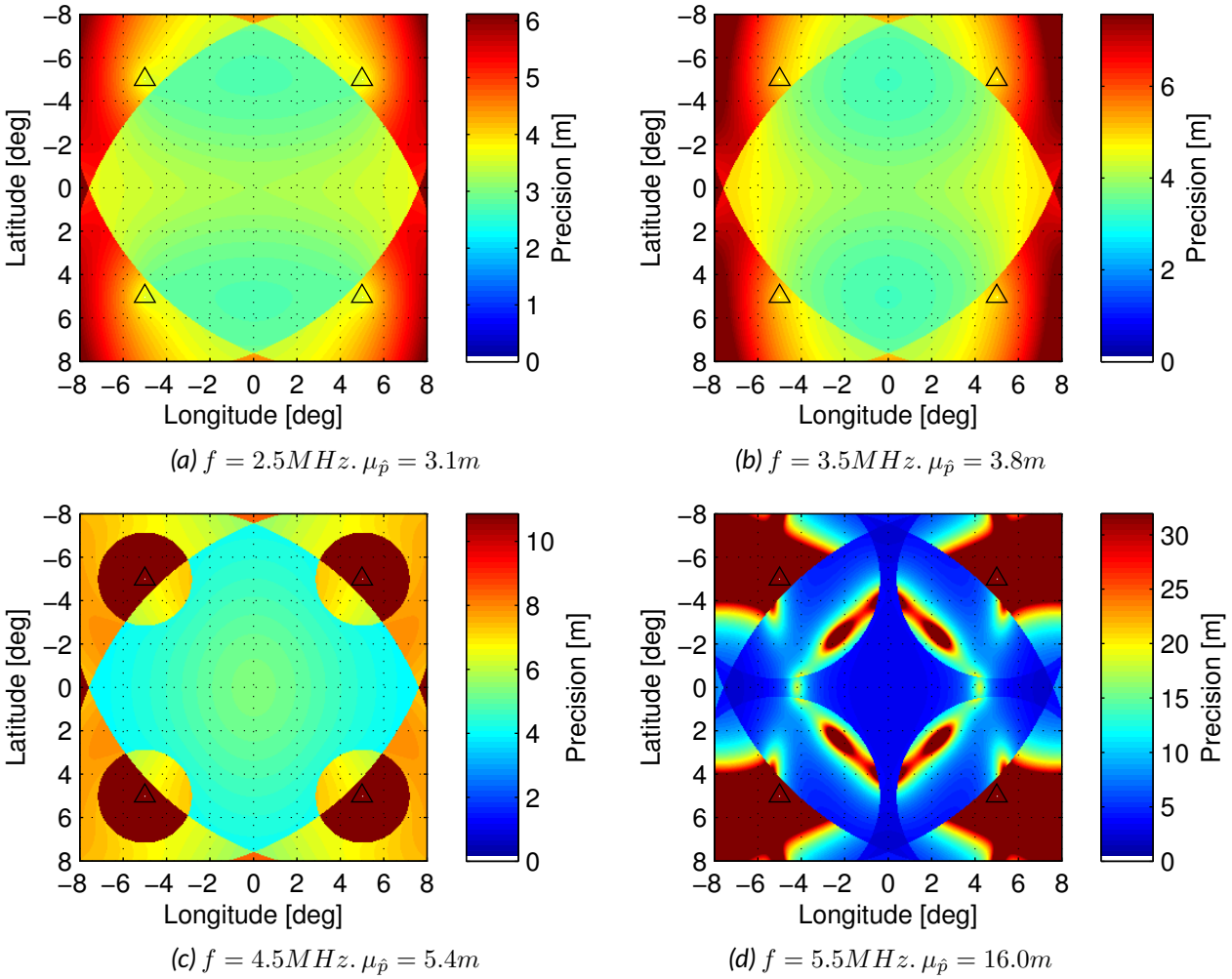
$$\boldsymbol{M_P}(f = 2.5MHz) = diag(2.22 \cdot 10^9, 2.22 \cdot 10^9, 2.22 \cdot 10^9, 2.22 \cdot 10^9)\ s^{-2} \quad (10.5)$$

$$\boldsymbol{M_P}(f = 3.5MHz) = diag(1.69 \cdot 10^9, 1.69 \cdot 10^9, 1.69 \cdot 10^9, 1.69 \cdot 10^9)\ s^{-2} \quad (10.6)$$

$$\boldsymbol{M_P}(f = 4.5MHz) = diag(6.94 \cdot 10^8, 6.94 \cdot 10^8, 6.94 \cdot 10^8, 6.94 \cdot 10^8)\ s^{-2} \quad (10.7)$$

$$\boldsymbol{M_P}(f = 5.5MHz) = diag(3.35 \cdot 10^9, 3.35 \cdot 10^9, 3.35 \cdot 10^9, 3.35 \cdot 10^9)\ s^{-2} \quad (10.8)$$

A solution for the precision can be found with the use of equation 10.3 and 10.4, and the ionospheric composition from the nominal scenario, defined in section 6.1. Inserting equa-



(a) $f = 2.5MHz.\ \mu_{\hat{p}} = 3.1m$

(b) $f = 3.5MHz.\ \mu_{\hat{p}} = 3.8m$

(c) $f = 4.5MHz.\ \mu_{\hat{p}} = 5.4m$

(d) $f = 5.5MHz.\ \mu_{\hat{p}} = 16.0m$

Figure 10.2: The obtained performance for a TOA -based positioning system for four frequencies: fig. 10.2a $2.5MHz$, 10.2b $3.5MHz$, 10.2c $4.5MHz$, 10.2d $5.5MHz$. The setup of the positioning system is such that four beacons are placed in the corners at latitudes $\beta = \pm 5.0°$ and $\lambda = \pm 5.0°$, indicated by triangles in the figures. The subsolar point is at (0,0). Performance is evaluated by a search over all possible locations of the mobile station, yielding the heatmaps above. Note that the scale of the colors is defined differently for each subfigure.

tions 10.3 and 10.4 into equation 9.9 yields a solution for the Jacobian, which in turn can be used to compute the Fisher information matrix and subsequently the CRLB , ultimately leading to the estimated precision. To obtain a quantitative result, the values for the least-squares fit from table 10.1 are used. The result of this endeavor is a heatmap showing the performance of an ideal TOA -based positioning system under pre-specified conditions, shown in figure 10.2.

The performance is evaluated for the nominal scenario, c.f. table 6.1, and for four independent system frequencies. A comparison of figures 10.2a to 10.2d show that for all cases, the best performance is obtained centrally between the beacons. However, in the analysis, the beacons are placed in a perfect grid. In reality this might prove to be impossible due to terrain constraints, negatively influencing the obtained performance. Furthermore it has been assumed that positioning is done on a spherical planet. Hence, fluctuations in the terrain are not taken into account. The resulting performance is thus only an ideal case.

The initial results from figure 10.2 in an ideal setup also demonstrate that in all cases, a system with a lower system frequency yields better performance results. This can be explained by considering the increase in standard deviation with higher frequencies, as shown in table 10.1. Furthermore, the skip distance has a profound effect on the performance. At areas around the beacons, a jump in performance can be seen. This is caused by a loss of signal from the nearby beacon, degrading the availability to only three beacons, causing a performance degradation.

10.2 Angle of Arrival

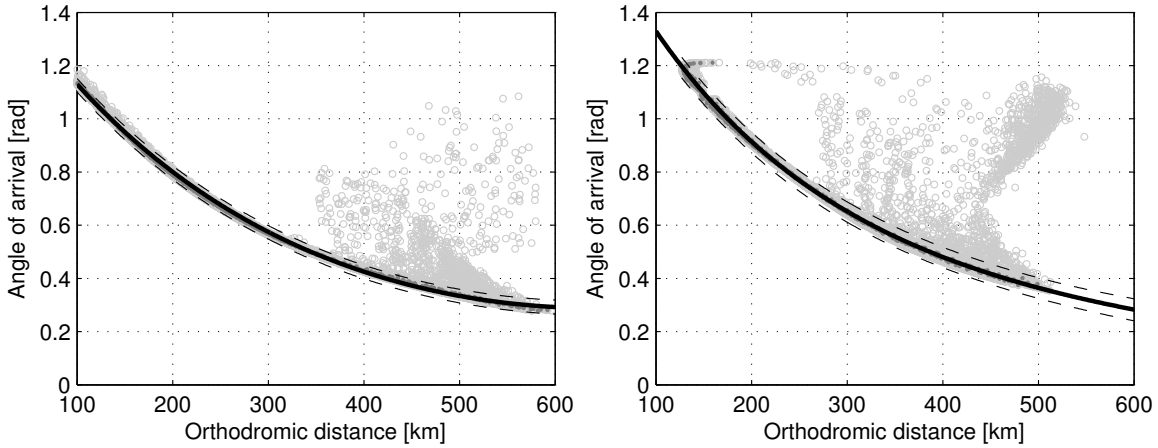
A receiver using an AoA -based approach determines its position by interpreting the incident elevation angle of signals from different beacons. In order to resolve a position estimate in 2D-space, only two signals from two independent beacons are required. However, an obvious complication is the nature of the NLOS ionospheric environment in which the system has to operate. Signals are refracted differently depending on their launch elevation and initial location, causing diffusion and distortion in the angle of arrival. This section treats a similar approach to obtaining the CRLB as was done wit a TOA -based approach. However, this time the angle of arrival is used.

In order to find the CRLB for an AoA -based approach, it is necessary to measure the angle of arrival of a signal from a certain beacon, and relate it to the distance to that beacon. As such, the measurement parameter is the angle of arrival, denoted γ_a. The simulation has been used to find the relationship between this angle and the distance, as is shown in figure 10.3. The same outlier rejection method is used as with the TOA method. Data points which are rejected are depicted light-gray in the figure, whereas data points which are accepted are dark gray. Note that the same dataset is used as for the TOA approach. As a result, the same skip distances are found for the AoA approach, which is evident from the figure.

In order to find a relation between the orthodromic distance and the angle of arrival, an analytical expression is fitted to the data. The best fit turns out to be exponential, yielding the following relationship:

$$\gamma_a = a\exp(bd) + c\exp(gd) \tag{10.9}$$

The constants a, b, c and g depend on the current state of the environment as well as the ray frequency and launch azimuth. The distance to a beacon is indicated by d. The results in figure 10.3 are specifically obtained for a ray frequency of $2.5MHz$ and $4.5MHz$, in figures



(a) $f = 2.5MHz, \bar{\sigma} = 0.47deg$ *(b)* $f = 4.5MHz, \bar{\sigma} = 0.70deg$

Figure 10.3: A least-squares exponential fit of the angle of arrival as a function of the orthodromic distance for four frequencies. Solid black lines indicate the fit and its 95% *confidence levels (dashed). Points which were used for the fit are depicted in dark gray whereas points which were discarded are depicted in light gray. In both cases, a beacon position of* $(0^\circ, 0^\circ)$ *in lat/lon coordinates has been used, with an azimuth of* 0°.

Table 10.2: Model fit parameters for various frequencies where AoA is used. The exponential fit has the shape $\gamma_a = a\exp(bd) + c\exp(gd)$. *The data is obtained using the nominal scenario.*

f	$2.5MHz$	$3.5MHz$	$4.5MHz$	$5.5MHz$
a $[rad]$	1.594	1.617	1.227	7.09
b $[km^{-1}]$	$-3.558 \cdot 10^{-3}$	$-3.241 \cdot 10^{-3}$	$-5.803 \cdot 10^{-3}$	$-1.68 \cdot 10^{-2}$
c $[rad]$	$5.993 \cdot 10^{-3}$	$1.180 \cdot 10^{-5}$	0.772	1.642
g $[km^{-1}]$	$4.751 \cdot 10^{-3}$	$1.528 \cdot 10^{-2}$	$-1.912 \cdot 10^{-3}$	$-2.893 \cdot 10^{-3}$
1σ	$8.20 \cdot 10^{-3} rad$	$1.46 \cdot 10^{-2} rad$	$1.23 \cdot 10^{-2} rad$	$2.44 \cdot 10^{-2} rad$

10.3a and 10.3b respectively. Table 10.2 gives an overview of the expression constants for various frequencies.

The next step is to construct and derive the measurement matrix for the angle of arrival. This matrix is constructed for the number of beacons n:

$$\gamma_{\mathbf{a}} = \begin{bmatrix} a\exp(bd_1) + c\exp(gd_1) \\ a\exp(bd_2) + c\exp(gd_2) \\ \vdots \\ a\exp(bd_n) + c\exp(gd_n) \end{bmatrix} \tag{10.10}$$

The Jacobian of the measurement matrix, defined in equation 10.10, is obtained by first combining it with the expression for the orthodromic distance, previously defined in equation 8.2. Subsequently, the matrix can be derived with respect to both the longitude and latitude. The details of this derivation can be found in appendix A.3, equations A.23 and

A.25:

$$\frac{\partial \gamma_a}{\partial \beta} = \left[ab\exp^{bd_n} + cg\exp^{gd_n}\right]\frac{-R_M}{\sqrt{1-x^2}} \cdot (\cos(\beta_{BC})\sin(\beta_{MS})\cos(\Delta\lambda) + \sin(\beta_{MS})\cos(\beta_{BC})) \quad (10.11)$$

$$\frac{\partial \gamma_a}{\partial \lambda} = \left[ab\exp^{bd_n} + cg\exp^{gd_n}\right]\frac{R_M}{\sqrt{1-x^2}} \cos(\beta_{BC})\cos(\beta_{MS})\sin(\Delta\lambda) \quad (10.12)$$

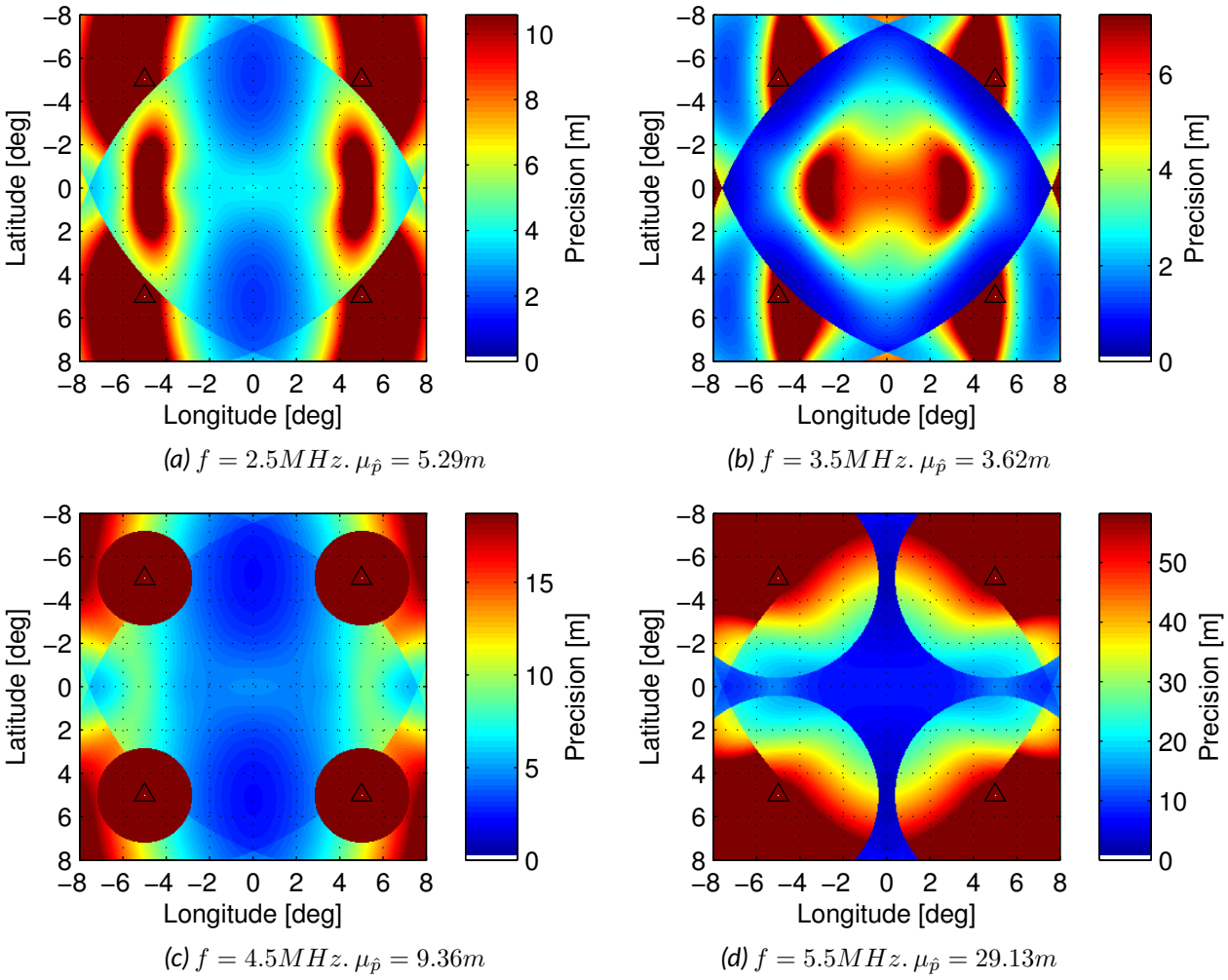


Figure 10.4: The estimated performance for an AoA -based beacon system. The performance map is obtained with four beacons, positioned at latitudes and longitudes $\pm 5°$. The subsolar point is at $(0°,0°)$. Note that the scale of the colors is defined differently for each subfigure. An unambiguous result from the figures is that the performance decreases with increasing frequency.

The measure of precision $\boldsymbol{M_P}$ is again found to consist of standard deviations are equal, as it was found that they are equal within 1 %. Thus, the $\boldsymbol{M_P}$ for the AoA -approach equals for the simulated frequencies:

$$\boldsymbol{M_P}(f = 2.5MHz) = diag(1.49 \cdot 10^4, 1.49 \cdot 10^4, 1.49 \cdot 10^4, 1.49 \cdot 10^4)\, rad^{-2} \quad (10.13)$$

$$\boldsymbol{M_P}(f = 3.5MHz) = diag(4.69 \cdot 10^3, 4.69 \cdot 10^3, 4.69 \cdot 10^3, 4.69 \cdot 10^3)\, rad^{-2} \quad (10.14)$$

$$\boldsymbol{M_P}(f = 4.5MHz) = diag(6.61 \cdot 10^3, 6.61 \cdot 10^3, 6.61 \cdot 10^3, 6.61 \cdot 10^3)\, rad^{-2} \quad (10.15)$$

$$\boldsymbol{M_P}(f = 5.5MHz) = diag(1.68 \cdot 10^3, 1.68 \cdot 10^3, 1.68 \cdot 10^3, 1.68 \cdot 10^3)\, rad^{-2} \quad (10.16)$$

Like with the TOA -approach, the derivations for the latitude and longitude, obtained in equations 10.11 and 10.12 respectively, are used to compute the CRLB . The computations use the values for the model constants, found in table 10.2. The performance of an AoA -based system is then calculated for various locations of the mobile stations and for various frequencies. This leads to the heatmap shown in figure 10.4. In this figure, the performance is shown for a nominal scenario. It is evident from the individual maps that the performance degrades as the system frequency increases. The skip zone is indistinguishable by the circular areas around the beacons on the map. In these areas, the performance takes a sudden drop, caused by the loss of signal from the nearby beacon. This loss of signal is a direct result of the skip distance. As the frequency increases, the skip distance increases, an effect which is reflected as well by the increase of the circular areas around the beacons. The mean precision increases particularly fast for system frequencies above the plasma frequency, which is $4.45MHz$ in this case.

10.3 Received Signal Strength

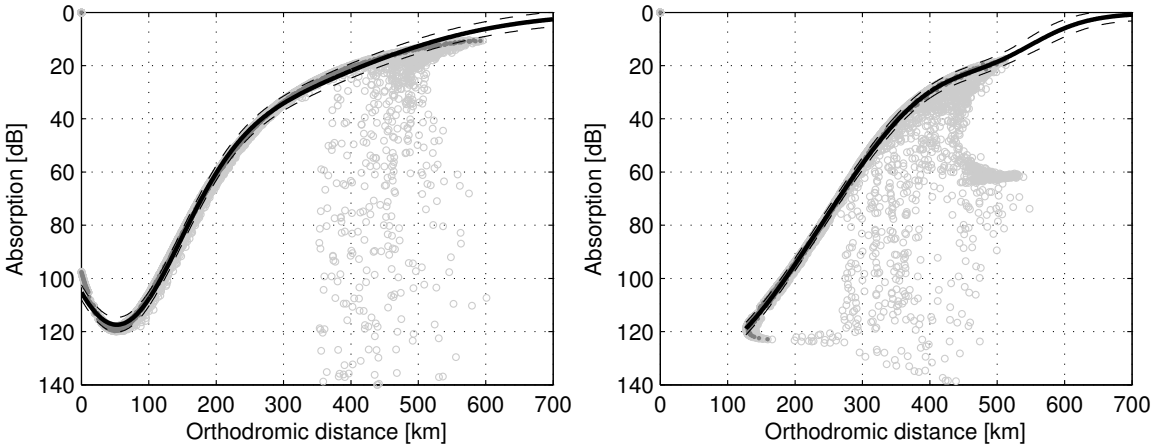
In a RSS -based approach, the signal strength is measured and related to the distance through some predefined model fit. The signal strength in an NLOS environment does not decrease with distance linearly. Instead, it peaks at a certain distance from a beacon. Clearly, the received signal strength, denoted by P_r, is the measurement parameter in this approach. In this section, the model fit is obtained by the use of simulation data. Subsequently, the precision is calculated as has been done in sections 10.1 and 10.2.

A result of the simulated signal strength as a function of the distance is shown in figure 10.5. In here, the skip distance is evident from the lack of signal strength at shorter distances. Furthermore, the signal strength peaks at a certain distance away from the beacon. This is in line with the theory from section 8.2.1, where the effects of an NLOS environment on the signal strength have been described.

An analytical expression can be derived from the measurements in figure 10.5. A second-order Gaussion fit of the measurements is expressed as follows:

$$P_r = a_1 \exp\left(-\left(\frac{d_n + b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{d_n + b_2}{c_2}\right)^2\right) \quad (10.17)$$

The six constants in equation 10.17 vary with frequency and ionospheric composition. Hence, evaluation for variations in these external variables is essential. Table 10.3 expresses the constants found for variations in system frequency. In this table, the standard deviation of the absorption is also given for each system frequency. It is expected that the precision of



(a) $f = 2.5MHz, \bar{\sigma} = 1.39dB$ *(b)* $f = 4.5MHz, \bar{\sigma} = 1.20dB$

Figure 10.5: A least-squares fit of the received signal strength, expressed as absorption, as a function of the orthodromic distance for various frequencies. Solid black lines indicate the fit and its 95% *confidence levels (dashed). Points which were used for the fit are depicted in dark gray, whereas points which were discarded are depicted in light gray. In both cases, a beacon position of* $(0°, 0°)$ *in lat/lon coordinates has been used, with an azimuth of* $0°$.

the whole system will decrease with increasing frequency, as the standard deviation of the absorption rises with increasing frequency.

The measurement matrix for n beacons is obtained from the definition of the second-order gaussian fit, stated in equation 10.17:

$$\mathbf{P_r} = \begin{bmatrix} a_1 \exp\left(-\left(\frac{d_1+b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{d_1+b_2}{c_2}\right)^2\right) \\ a_1 \exp\left(-\left(\frac{d_2+b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{d_2+b_2}{c_2}\right)^2\right) \\ \vdots \\ a_1 \exp\left(-\left(\frac{d_n+b_1}{c_1}\right)^2\right) + a_2 \exp\left(-\left(\frac{d_n+b_2}{c_2}\right)^2\right) \end{bmatrix} \tag{10.18}$$

The derivation of equation 10.18 yields the Jacobian for the RSS method:

$$\begin{aligned} \frac{\delta \mathbf{P_r}}{\delta \beta} &= \left[\frac{(d_n - b_1)a_1}{c_1} \exp\left(-\left(\frac{d_n - b_1}{c_1}\right)^2\right) + \frac{(d_n - b_2)a_2}{c_2} \exp\left(-\left(\frac{d_n - b_2}{c_2}\right)^2\right)\right] \\ &\cdot \frac{2R_M}{\sqrt{1-x^2}} \left[\cos(\beta_{BC}) \sin(\beta_{MS}) \cos(\Delta\lambda) + \sin(\beta_{BC}) \cos(\beta_{MS})\right] \end{aligned} \tag{10.19}$$

$$\begin{aligned} \frac{\delta \mathbf{P_r}}{\delta \lambda} &= \left[\frac{(d_n - b_1)a_1}{c_1} \exp\left(-\left(\frac{d_n - b_1}{c_1}\right)^2\right) + \frac{(d_n - b_2)a_2}{c_2} \exp\left(-\left(\frac{d_n - b_2}{c_2}\right)^2\right)\right] \\ &\cdot \frac{2R_M \cos(\Delta\lambda - 1)}{\sqrt{1-x^2}} \left[\cos(\beta_{BC}) \cos(\beta_{MS}) \sin(\Delta\lambda)\right] \end{aligned} \tag{10.20}$$

For the RSS method, it has again been found that the standard deviations are equal within

1 %. The measure of precision yields:

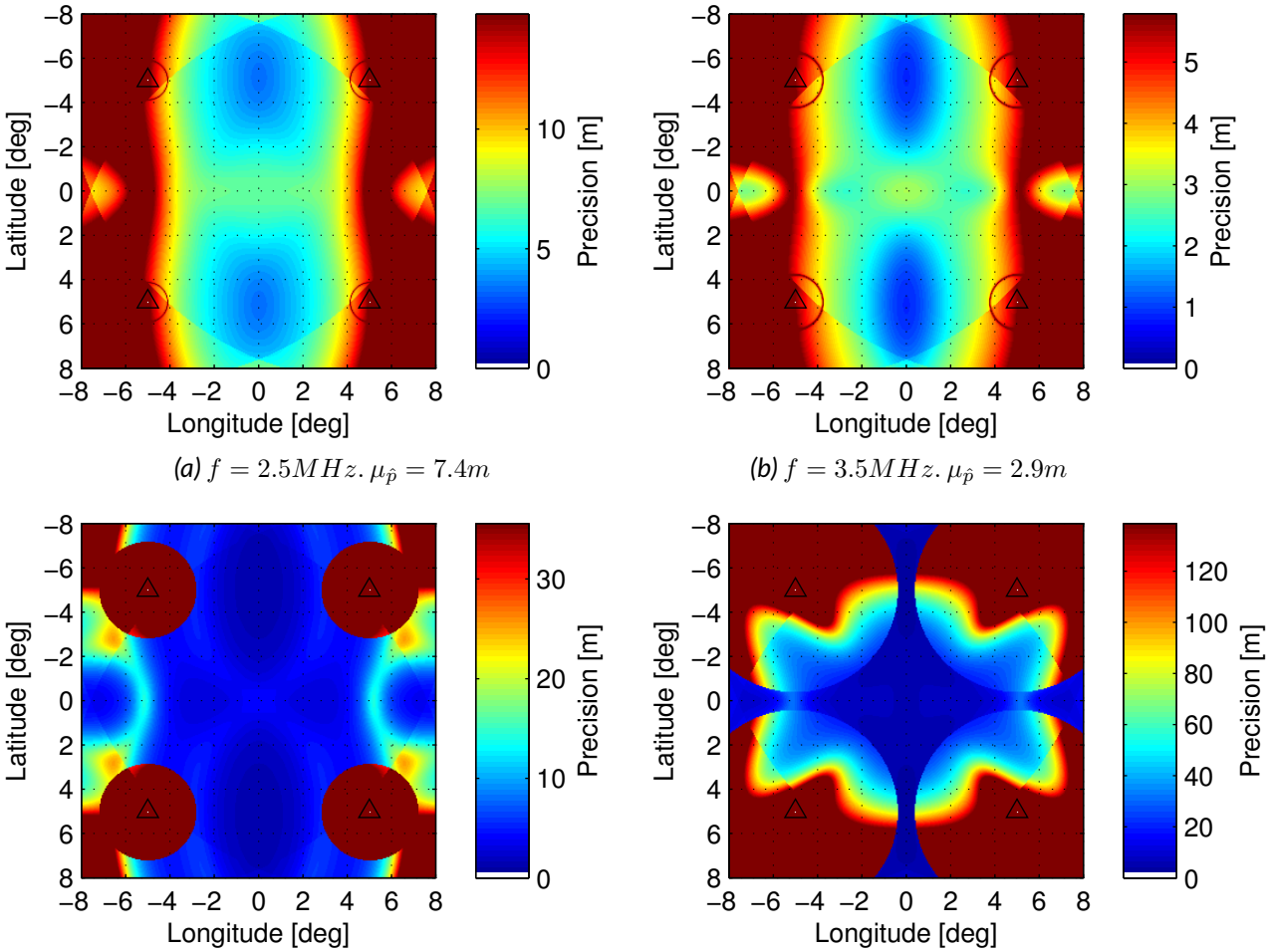
$$\boldsymbol{M_P}(f = 2.5MHz) = diag(1.386^{-2}, 1.386^{-2}, 1.386^{-2}, 1.386^{-2})\, dB^{-2} \quad (10.21)$$

$$\boldsymbol{M_P}(f = 3.5MHz) = diag(0.613^{-2}, 0.613^{-2}, 0.613^{-2}, 0.613^{-2})\, dB^{-2} \quad (10.22)$$

$$\boldsymbol{M_P}(f = 4.5MHz) = diag(1.195^{-2}, 1.195^{-2}, 1.195^{-2}, 1.195^{-2})\, dB^{-2} \quad (10.23)$$

$$\boldsymbol{M_P}(f = 5.5MHz) = diag(3.063^{-2}, 3.063^{-2}, 3.063^{-2}, 3.063^{-2})\, dB^{-2} \quad (10.24)$$

The full derivations of equations 10.19 and 10.20 can be found in appendix A.2. The derivations obtained in equations 10.19 are used to compute the CRLB for the RSS -method. The results are shown in the heatmaps in figure 10.6 for various frequencies. Note that the precision does not decrease linearly with increasing frequency. The best precision is obtained for a frequency of $3.5MHz$. Interestingly, the RSS -based approach outperforms all other approaches in terms mean precision. The RSS -based approach with system frequency of $3.5MHz$ shows the best performance of all technologies treated so far. There



(a) $f = 2.5MHz.\ \mu_{\hat{p}} = 7.4m$

(b) $f = 3.5MHz.\ \mu_{\hat{p}} = 2.9m$

(c) $f = 4.5MHz.\ \mu_{\hat{p}} = 21.1m$

(d) $f = 5.5MHz.\ \mu_{\hat{p}} = 68.3m$

Figure 10.6: The estimated performance for a RSS -based beacon system. The performance map is obtained with four beacons, positioned at latitudes and longitudes $\pm 5°$. The subsolar point is at ($0°$,$0°$). Note that the scale of the colors is defined differently for each subfigure.

Table 10.3: Model fit parameters for various frequencies where RSS is used. The model fit is defined in equation 10.17. The data is obtained using the nominal scenario.

f	$2.5MHz$	$3.5MHz$	$4.5MHz$	$5.5MHz$
a_1 $[dB]$	-76.83	-104.1	-132.9	-11.37
b_1 $[m]$	47.49	65.14	28.98	264.4
c_1 $[m^{-1}]$	132.8	156.5	294.1	43.11
a_2 $[dB]$	-42.16	32.52	-8.36	-68.77
b_2 $[m]$	118.1	211.1	500.5	225.0
c_2 $[m^{-1}]$	348.3	337.9	96.11	242.5
1σ	$1.386dB$	$0.6130dB$	$1.195dB$	$3.063dB$

is a significant discrepancy in performance for system frequencies above the plasma frequency ($4.45MHz$), compared to those below. This discrepancy is most evident from the RSS -based approach, but the data for the other approaches suggest a similar discrepancy. Note the difference in mean precision for various frequencies in figure 10.2 and 10.4.

10.4 Summary

The results from the previous sections have shown that the best results are obtained by using a RSS -based approach. Both the AoA -based and TOA -based approach yield comparable results. Nevertheless, the performance of all approaches is comparable. Furthermore, a higher ray frequency is harmful for all technologies. It is evident that an increase in frequency causes a decrease in precision, no matter which technology is selected. Specifically, a significant degrade in performance is seen for system frequencies above the plasma frequency. Therefore, a sound conclusion is to assure that the used frequency stays below the plasma frequency.

So far, a comparison has only been made under a nominal scenario and ideal conditions. The question is how the precision responds with a varying environment. The next chapter therefore focuses on exploring the performance with different scenario's and different beacon placements.

11 — Performance Envelope

The overall performance of each technology is determined by studying their respective performance envelope. So far the performance of various technologies has been evaluated for nominal ionospheric conditions. The performance envelope is defined as a set of limits within which some localization technology can provide a position estimate.

Several factors generally limit a ground-based localization system like the one in this work. First, there is the ionosphere, which continuously changes. Second, the frequency affects the propagation of signals, causing variations in performance. Third, in some cases, the reception of beacon signals will be too weak to determine a position at all. These events will limit the area in which a position estimate can be obtained, and thus ultimately the coverage that a localization system provides. Last, there is the issue of beacon placement. It is expected that adding more beacons to the system will increase performance by enhancing the coverage and precision. Adding more beacons might mitigate the impact of the skip distance on the precision, which was previously discovered in chapter 10.

In this chapter, the system parameters are varied to find operational limits and optimal characteristics. Previously, scenarios have been defined which classify the state of the ionosphere. Section 11.1 demonstrates the impact of the variations in the ionosphere under these scenarios. Next, frequency effects are explored in section 11.2. The technical limitations, imposed by the antennae used, are explored in section 11.3. Section 11.4 examines the effects and optimal coverage of the locations of individual beacons. Finally, section 11.5 gives an overview of the results found in this chapter.

11.1 Ionospheric effects

Different performance results are obtained with different ionospheric compositions. Recall that the ionospheric compositions have been previously defined in *scenarios* in section 6.1, specifically in table 6.1. This section analyses the performance results for various scenarios by comparing the mean precision that can be obtained in each scenario. The following five scenarios are evaluated:

- *Nominal*: the nominal scenario
- *Max*: maximum Solar interaction
- *Min*: minimum Solar interaction
- *Duststorm*: an ionosphere during a dust storm
- *SEP*: an ionosphere during a SEP event

The results for each scenario are shown in figure 11.1. The figure demonstrates that the best performances can be obtained with both the TOA and AoA -approach. Using these approaches, a sub-10 meter resolution is achieved in practically all circumstances. The results for the RSS -approach are not so unambiguous. For most scenarios, the RSS -approach yields performance comparable to the other approaches. However, dissimilar results are obtained for the *min* and *SEP* scenarios. Conclusively, the RSS approach deems to be not reliable under all circumstances. The TOA and AoA approaches are stable, yielding a position estimate under any scenario. According to figure 11.1, the TOA approach achieves a preci-

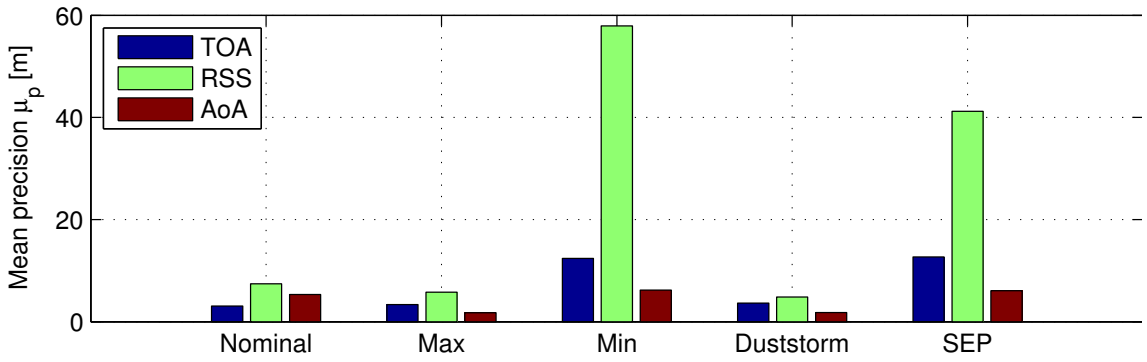
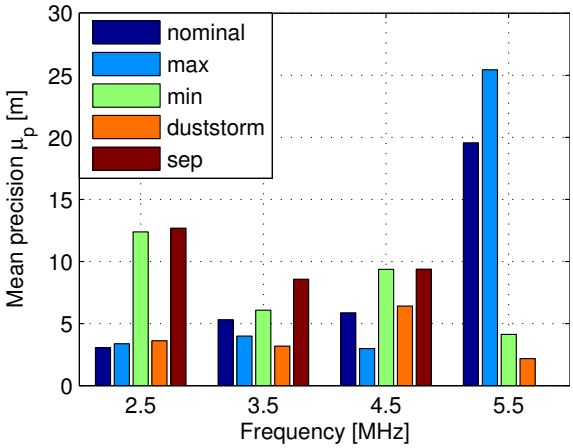
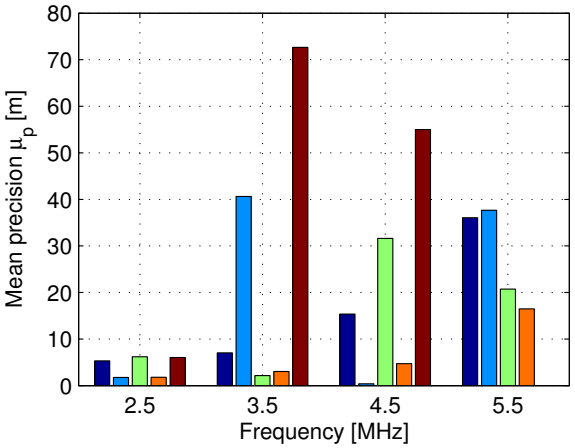


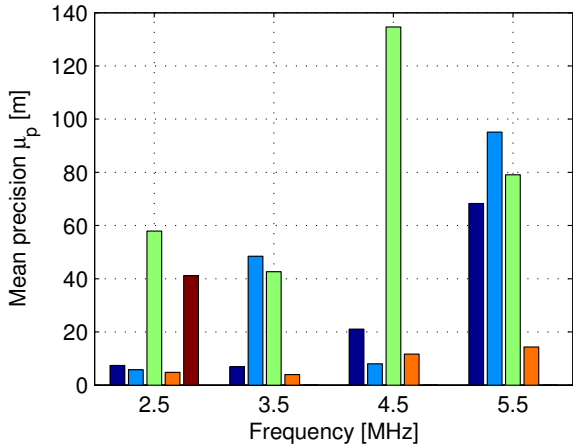
Figure 11.1: A comparison of the obtained performance using a TOA, RSS or AoA approach for all scenarios. The operating frequency is $2.5MHz$. The measures of precision $\boldsymbol{M_P}$, used to obtain these results, are found in appendix B.



(a) Performance of the TOA approach



(b) Performance of the AoA approach



(c) Performance of the RSS approach

Figure 11.2: An overview of the performance at operating frequencies from $2.5MHz$ to $5.5MHz$, with steps of $1MHz$. Some results are missing as the system was unable to yield any position estimate under those conditions. The measures of precision $\boldsymbol{M_P}$, used to obtain these results, are found in appendix B.

sion of less than $13m$ under any scenario, when a frequency of $2.5MHz$ is used. Specifically, the TOA -approach attains a precision of less than $8m$ for the *nominal* scenario. For the same frequency, the AoA -approach achieves a precision less than $6m$, and the RSS -approach a precision less than $58m$ for any scenario. Clearly, the AoA -approach outperforms the others in terms of minimum obtainable precision for the aforementioned frequency. A study on frequency variations can determine if this conclusion holds for all operational frequencies.

11.2 Frequency effects

It is found that the operating frequency heavily impacts the performance of a positioning system. The performance has been evaluated for various operating frequencies. This is done for all approaches, and for all scenarios.

The performances as a function of frequency are depicted in figure 11.2. The results are consistent with the conclusion obtained in section 11.1. The best performance is obtained by using either the TOA approach, shown in figure 11.2a, or the AoA -approach, shown in figure 11.2b. Although the AoA approach obtains the most consistent results for all scenarios at a frequency of $2.5MHz$, this approach shows more fluctuations at higher frequencies. The results are more inconsistent at higher frequencies. Especially, the performance under the *SEP* scenario is poor, as it quickly degrades with increasing frequencies.

The performance is worse across the board when the RSS approach is used. Still, the best performance for this method is obtained at an operating frequency of $2.5MHz$. This outcome is consistent with that of the performance under the TOA approach. The largest discrepancies in performance for different scenarios are seen for frequencies of $4.5MHz$ and $5.5MHz$. For lower frequencies, the discrepancies are smaller, a result which is consistent with that for the TOA approach.

The frequency analysis shows that a lower frequency is beneficial in terms of performance, stability and consistency. The best performance overall is obtained with the AoA -approach under a frequency of $2.5MHz$, yielding an upper bound on the precision of $6m$. This metric entails that, for any scenario, the precision will always less than the upper bound. An overview of the upper bounds of all combinations of frequency and technology is given in table 11.1. Although the AoA -approach may yield the best upper bound, it is evident from table 11.1 that the TOA -approachs has a more favorable upper bound for higher frequencies. The upper bound never raises above $26m$. Therefore, it is said that the TOA approach is the most stable. The RSS -approach underperforms significantly.

Table 11.1: A comparison of the upper bound *on the precision for a given frequency and technology. The upper bound on the precision is the precision with which an estimate can be obtained for any scenario. The same results are also evident from figure 11.2.*

Frequency	TOA	AoA	RSS
2.5 MHz	$13m$	$6m$	$58m$
3.5 MHz	$8m$	$73m$	$47m$[1]
4.5 MHz	$9m$	$55m$	$138m$[1]
5.5 MHz	$26m$[2]	$38m$[2]	$95m$[1]

[1] No position estimation was possible for the *SEP* and *min* scenario.
[2] No position estimation was possible for the *SEP* scenario.

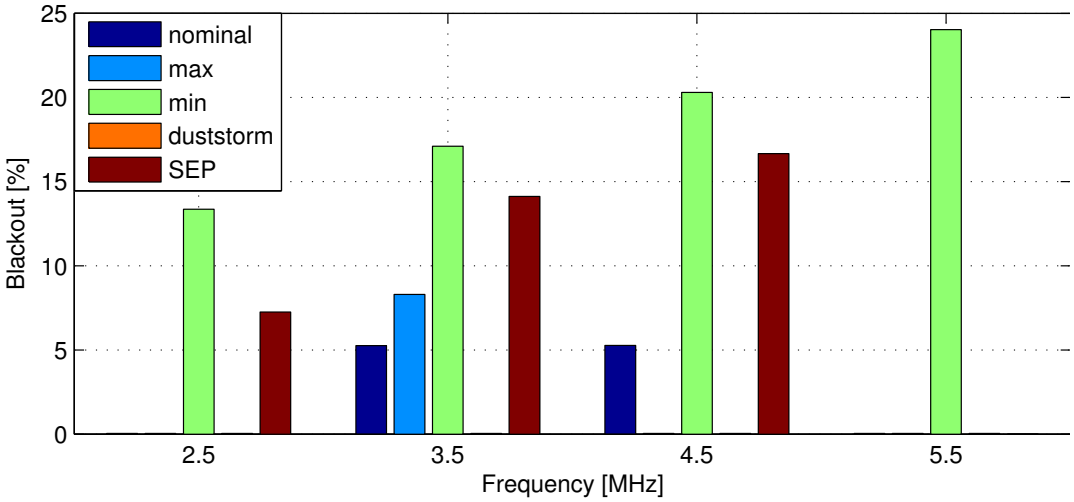


Figure 11.3: An overview of the blackout factor, expressed in percentages, of all localization technologies for operating frequencies from $2.5MHz$ to $5.5MHz$, with steps of $1MHz$. Under some circumstances, practically no blackout is experienced. The measures of precision $\boldsymbol{M_P}$, used to obtain these results, are found in appendix B.

11.3 Antenna effects

Signals with too little signal strength can not be properly received by a mobile station, causing a loss of signal and degrading the precision of a performance estimate. In chapter 7, it was determined that a maximum absorption of $119.64dB$ is tolerable, regardless of elevation angle. The absorption is calculated in equation 4.38 for the entire path of the ray, and then compared to this lower bound. Any signal below this lower bound is lost and cannot be used for position estimation.

The *blackout* factor, Bf, represents the percentage of signals which arrive at a mobile station, but with a signal strength below the lower bound. The blackout factor for one signal from one beacon is binary; a signal is either received successfully or it is not. In a situation where n beacons are available, the number of beacons from which a signal is received successfully is indicated by n_s. Thus, when four beacons are present, the blackout factor is $1 - n_s/n = \{0, 0.25, 0.5, 0.75, 1\}$. Note that blackout is the opposite of coverage. In an area where a blackout persists, no coverage can be provided.

The blackout factor in itself varies with frequency and distance. The *mean blackout factor* expresses the average factor of blackout, experienced by a mobile station, over a certain area between a set of beacons. It is a measure which indicates how severe the performance estimate is degraded due to a lack of a successfully received signal. The mean blackout signal is expressed as follows:

$$\mu_{Bf} = \frac{1}{N^2} \sum_{\beta=1}^{N} \sum_{\lambda=1}^{N} 1 - \frac{n_{s,\beta,\lambda}}{n_{\beta,\lambda}} \tag{11.1}$$

In equation 11.1, the area, in which a mobile system operates, is spanned by a certain range in longitude λ and latitude β, evaluated in N steps.

The mean blackout factor depends on the chosen scenario and frequency. Figure 11.3

gives an overview of the blackout factor for various scenarios and frequencies. The results indicate that blackout is not a significant issue for the *nominal*, *max* and *duststorm* scenario. Higher percentages are encountered with the *min* and *SEP* scenarios. The blackout factor is seen steadily rising with frequency. This behavior equals the results found in section 6.4.2, where it was shown that an increase in frequency causes an increase in absorption. This, in turn, causes more blackout.

It is evident from figure 11.3, that the lowest blackout factors are obtained at a frequency of $2.5MHz$. The blackout factor escalates with rising frequency. Hence, from the point of successful signal reception, it is concluded that a lower frequency is beneficial. All in all, the coverage provided by the system is high, as the blackout factors for the *nominal* and *max* scenarios is below 8 %, as is evident from figure 11.3. Hence, the coverage is above 92 % for these scenarios. In the case of a *min* scenario, the blackout is 13 % for an operating frequency of $2.5MHz$, and it steadily rises for higher frequencies. The *SEP* scenario shows similar results. Curiously there is no blackout experienced at all for the duststorm scenario. One possible explanation for this could be the fact that the ionospheric peak density is lifted during a dust storm. Radio signals are refracted more before reaching the peak, hence experiencing a lower total absorption.

11.4 Location effects

The location affects the performance in two ways. First, a signal can be lost due to absorption if a beacon is placed too far away. Second, the geometric dilution of precision may be too high, which makes that a position estimate must be discarded.

11.4.1 Beacon placement

The placement of the beacons influence the quality of signals received at the mobile station, and consequently, the obtained performance for a position estimate. Several factors limit an optional placement of the beacons. First, the skip distance limits reception close to a beacon, as was proven in section 6.3.2. Second, the models obtained in chapter 10 are only valid in the range for which a model fit could be made from the simulation data. This range is lower bounded by the skip distance, and upper bounded by some maximum distance c.f. figure 6.8. Third, the limits on the signal strength dictate when a signal is lost due to low signal strength. Section 6.4.2 proved that the absorption increases with distance. Hence, an upper bound exists, defined by the maximum allowed absorption. All these bounds themselves depend on many factors, like the ionospheric composition and the operational frequency which is used.

Furthermore, the previous sections have shown that the performance can be poor for certain scenarios or localization technologies. This leads to the question if beacons can be placed such, that the performance is enhanced under all scenarios. To examine this, several beacon arrangements are defined. An arrangement contains information about the location of beacons. So, each arrangement contains a certain number of beacons, placed at specific locations. It is expected that the arrangements with more beacons yield a better position estimate, as there are simply more beacons to provide positioning information. The following arrangements are defined:

- *RECT-S*: 4 beacons, placed in a square with sides of 5°
- *RECT-L*: 4 beacons, placed in a large square with sides of 10°

- *GRID-S*: 12 beacons, placed in a rectangular grid. Beacons are spaced by $10°$
- *TRIA-S*: 14 beacons, placed in a mesh of equilateral triangles with sides of $5°$
- *TRIA-L*: 14 beacons, placed in a mesh of equilateral triangles with sides of $10°$

A graphical representation of the five beacon arrangements is given in figure 11.4. The *RECT-L* arrangement is considered the standard; it has been used in this research so far.

A comparison of performance for different beacon arrangements is shown in figure 11.5. In the figure, only the performance for the *nominal* scenario (fig. 11.5a) and the *min* scenario (fig. 11.5b) is shown. It is clear that the beacon arrangements with more beacons yield a better performance. These are the *GRID-S*, *TRIA-S* and *TRIA-L* arrangements. The beacons in the *TRIA-S* arrangement, however, are packed closely together. As can be seen from figure 11.5, this only increases the performance marginally. In other words, the *GRID-S* and *TRIA-L* arrangement cover a lot more surface for a comparable performance to the *TRIA-S* arrangement. This is considered a downside for the *TRIA-S* arrangement, as a closely packed arrangement needs a lot more beacons to cover a certain area. When more beacons are needed, the mission cost rises. The mission complexity rises, too, as more beacons need

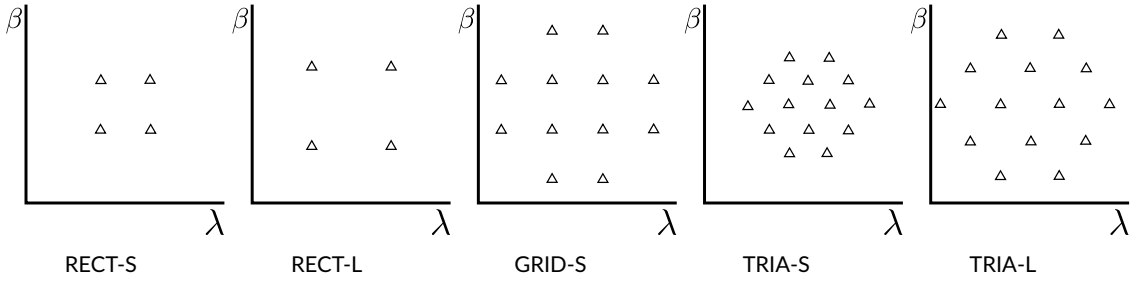
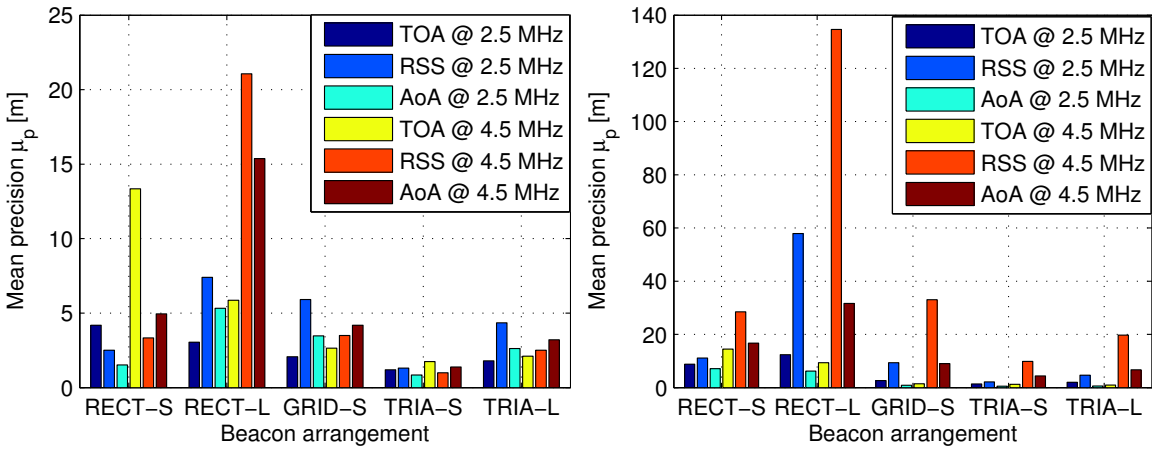


Figure 11.4: An illustration of the five beacon arrangements. Each triangle represents one beacon.



(a) Nominal *scenario* *(b)* Min *scenario*

Figure 11.5: An overview of the performance using a selection of different beacon arrangements. The operating frequencies are $2.5MHz$ and $4.5MHz$. The performance for the nominal scenario is shown in figure 11.5a, while figure 11.5b shows the performance under a minimal scenario. The measures of precision $\boldsymbol{M_P}$, used to obtain these results, are found in appendix B.

to be deployed. Therefore, the *GRID-S* and *TRIA-L* arrangements are superior to the *TRIA-S* arrangement.

The lowest geometric errors are obtained by using a *TRIA-S* beacon arrangement. The *GRID-S* and *TRIA-L* arrangements, compared in the figure, yield similar results. The *TRIA-S* arrangement provides a precision of at least $2m$ for the *nominal* scenario, and $10m$ for the *min* scenario, as is evident from figure 11.5. The *TRIA-L* arrangement, on the other hand, provides a precision below $4m$ and below $20m$, respectively, which is about twice as high as the *TRIA-S* arrangement. This precision can be obtained for both an operating frequency of $2.5MHz$ and $4.5MHz$. Discarding the RSS -approach, a precision below $10m$ is even possible. In the *min* scenario, both these arrangements and the *GRID-S* arrangement provide a sub-10-meter precision for all technologies, except the RSS -approach at an operating frequency of $4.5MHz$.

Considering that the mission cost and complexity is too high for the *TRIA-S* arrangement, it can be concluded that a precision of less than $4m$ and $10m$ are possible in the *nominal* and *min* scenario, respectively. Note that for the *nominal* scenario, this arrangement attains a precision twice as low for the TOA -approach.

The malfunction of beacons has so far not been considered. It is likely that a beacon will malfunction at some point in time. Hence, beacons should be placed such, that there are always three beacons within range of the mobile station. As was seen in section 8.1, three beacons are minimally required, and a fourth one is redundant. It is assumed that all arrangements contain a sufficient number of beacons to provide some level of redundancy. However, it is recommended to study the effects of beacon malfunction in more detail.

11.4.2 Dilution of precision

The geometric performance of a localization system is related to the geometry of the beacons in that system. To quantify the coverage, the GDOP is calculated for each technology, frequency and scenario. The GDOP has previously been explored in section 8.3. The GDOP value is a dimensionless variable, indicating the geometric error caused by the placement of the beacons in a system. It was found in section 8.3 that values around 1 indicate a low error, whereas GDOP values above 6 are generally discarded.

The GDOP has been evaluated for the *nominal* and *min* scenario in figure 11.6. Results for different frequencies and technologies have been grouped together per beacon arrangement. The standard deviations in the results are caused by this grouping. It was expected that the GDOP metric would be equal for any frequency or technology, as it is purely a geometric measure. However, it has been shown in section 11.3, that beacons are sometimes unavailable due to a weak signal strength. This poses an interesting problem, as the number of remaining beacons are reduced. Ultimately, a reduction in beacons increases the geometric error, and this is what causes the deviations in figure 11.6. Nevertheless, the deviations in GDOP are quite small compared to the results obtained for the beacon arrangements.

In all cases, however, the GDOP metrics show a good quality of the estimates. Refer to section 8.3 for a description of the numerical values. It remains to be seen what the impact is of mountainous terrain which can severely degrade the GDOP metric. This is left for future work.

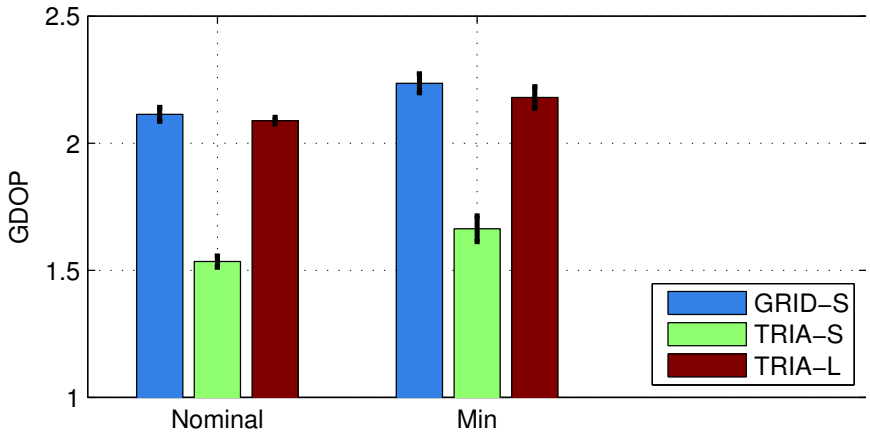


Figure 11.6: A comparison of GDOP results for the nominal *and* min *scenario. Error bars indicate the standard deviation of the GDOP caused by distinct frequencies and localization technologies.*

11.5 Summary

Several limitations persist on the use of a ground-based localization network on Mars.

First, ionospheric effects cause that some localization technologies cannot always be used. It is shown that the TOA and AoA -approaches are stable under any scenario, whereas the RSS -approach is not.

Second, the performance degrades unambiguously under an increase in frequency. The best results are obtained when lower frequencies are used. From the frequency analysis it is further apparent that the TOA -approach can yield a more stable performance than the AoA approach. Again, the performance using a RSS -approach is significantly worse than the other two.

Third, the antennas which are used limit the reception of signals, as some signals are too weak to be received. A *blackout factor* is used to measure the occurrences of loss of signal. Generally, an increase in frequency induces a higher fraction at which signals are lost. Furthermore, the *SEP* and *min* ionospheric scenarios are especially susceptible to blackout.

Last, the locations at which the beacons are placed have a profound effect on the performance of a localization system. The most optimal arrangements are the *GRID-S* and *TRIA-L* arrangements, where beacons are placed in a rectangular and triangular grid, respectively, and with distances of 10° between the beacons.

The results indicate that a precision of less than $4m$ can be obtained with the use of any technology in the *nominal* scenario with a *TRIA-S* arrangement, for frequencies of $2.5MHz$ and $4.5MHz$, Moreover, a precision of less than $10m$ for these frequencies is achieved, when the beacons are arranged in either an *GRID-S* or *TRIA-L* arrangement. using either an TOA or AoA approach. The *TRIA-S* arrangement is discarded, as it is found to be suboptimal from a mission cost & complexity perspective.

The limitations have shown that some approaches to a localization system on Mars are to be avoided, in order to guarantee a stable system with a good performance. Ultimately, a localization system using a TOA -approach with frequencies of 2.5 or $3.5MHz$, and a beacon arrangement with either rectangular or triagonally placed beacons, spaced 10° apart, yields the best and most stable performance.

12 — Conclusion & Recommendations

In this chapter, the conclusions to this work are presented in section 12.1. as this is a feasibility study, this research only explores the possibilities. Many possibilities, extensions and side-effects have not yet been accurately studied. Recommendations to the most important improvements are given in section12.2.

12.1 Conclusion

The goal of this work is to assert the feasibility of the use of a ground-based system, consisting of beacons, to aid in the provision of localization for mobile stations on Mars. It is shown that ground-based systems have inherent advantages to space-based systems. Accordingly, the hypothesis of this work is stated as follows:

A ground-based system can provide positioning information required by current and future Mars missions at a competitive precision, availability and coverage compared to a space-based positioning system.

A ground-based positioning system uses the ionosphere to refract radio signals. The Martian environment has been studied to determine how it inflicts with radio signals. During the dayside, a relatively stable ionosphere is found. The dayside ionosphere allows for radio signals to be refracted if their frequency is above the plasma frequency of an ionospheric layer, or reflected if their frequency is below the plasma frequency. This changes during the nightside, however, which is found to be very weak. The atmosphere of Mars is relatively faint, and hardly affects radio signals. Therefore, radio signals experience the most severe absorption in the ionosphere.

Mars does not have a strong global intrinsic magnetic field, but the ionosphere is surrounded by an induced magnetic field, caused by the Solar wind. Furthermore, the surface contains local crustal anomalies. However, due to the relatively low field strength, it has been assumed that the impact of the magnetic field on radio signals is insignificant.

A simulation tool is developed[1] to study the propagation of radio signals, which tracks the paths of radio signals through a simulated Martian environment. This simulation uses a Chapman-profile to model the ionosphere, which is laminated to numerically approximate the composition of an ionosphere. A ray-tracing method is used to track the signals. This method recursively traces radio signals, or *rays*, and changes their direction of propagation by having them interact with the environment. The result is a dataset which is used for analysis of positioning systems.

Several scenarios are defined to classify the state of the ionosphere. These scenarios are used in the simulation to obtain results for various extremes of ionospheric composition. A nominal scenario defines the ionospheric composition in its nominal state, having a peak electron density of $1.8 \cdot 10^{11} m^{-3}$ in the M2 layer. The *min* and *max* scenarios define an ionosphere at the point of minimum and maximum solar interaction, respectively, with a peak electron density in the M2 layer ranging from $1.4 \cdot 10^{11} m^{-3}$ to $2.6 \cdot 10^{11} m^{-3}$. Furthermore,

[1] The code for the simulation is available on *Github* (under an MIT license): `http://github.com/rvangijlswijk/ionospheric-ray-tracer`

a *dust storm* and *SEP* scenario are defined to account for dust storms and Solar energetic particle events. During a dust storm, the electron peak altitude rises to $137km$. The main impact of a SEP event is the rise of the peak electron density in the M1 layer, which rises to $1.2 \cdot 10^{11} m^{-3}$.

A comparison of the simulated dataset with existing measurements has shown that the simulation is able to model the Martian ionosphere with a high precision. Moreover, the *iris effect* emerges from the simulation results. This effect states that radio signals with an elevation angle above a critical angle are not reflected nor refracted. An increase in frequency causes an increase in the critical angle. The result set also indicates the existence of the skip distance above the plasma frequency, rising with increasing signal frequency.

The determination of the position of a mobile station can be done by exploiting the properties of the received signal. Three localization technologies are studied in detail in this work, being the *time of arrival* (TOA), *received signal strength* (RSS) and *angle of arrival* (AoA). The *Cramer-Rao lower bound* (CRLB) is used to determine the lower bound of the precision of a position estimate, produced by any of these technologies. The CRLB yields a performance metric for any given technology, provided a model to describe the measurements. This performance metric is expressed as the precision of a positioning system, measured in meters. In order to quantify the precision, a model of the relationship between the measured properties of the received signal and the position of the mobile station is to be determined. The simulation is used for this purpose. An analytical model is crafted from the simulation results for each of the three localization technologies. Subsequently, the precision of the position estimate of a mobile station is computed for various locations of the mobile station with respect to some beacons. A comparison of the mean precision for various frequencies shows that the the use of both the TOA and AoA technologies yields the best performance. Furthermore, an increase in frequency is found to be harmful to the obtained performance.

A comparison of the performance of all technologies in various scenarios shows that again, both the TOA and AoA -approach yield the most stable performance. Using these technologies, a precision of less than $6m$ for the AoA -approach, and a precision of less than $13m$ for the TOA -approach is obtained in any scenario. Furthermore, a coverage of at least 92 % is provided in all but the most extreme scenarios. Only in the *min* and *SEP* scenarios, a lower coverage is experienced. Here it is seen as well, that the coverage increases with decreasing frequency. The best results are always obtained with a frequency of $2.5MHz$.

The impact of the placement of beacons has been studied to determine the most optimal arrangement. It is shown that the best performance in the analyzed cases is obtained by placing beacons either in a rectangular or triangular grid on Mars, with beacons spaced 10° apart. By doing so, a precision of less than $4m$ can even be obtained under the nominal scenario for a frequency of $2.5MHz$ and $4.5MHz$ with any technology.

In conclusion, a ground-based beacon system is a competitive alternative to a space-based system. In the thesis motivation, it was shown current proposals for space-based systems indicate a precision of less than $10m$. The same precision can be achieved with a ground-based system, for a coverage of 92 %, at a lower mission cost & complexity than a space-based system.

12.2 Recommendations

This research has been done given an idealized situation. Several recommendations for further work are made to assert the feasibility of a ground-based beacon system in a wider set of conditions.

It has been assumed that Mars is a perfectly spherical planet, where variations in terrain have been ignored. Mountainous terrain, valleys and craters all will have a significant impact on the reception of radio signals. These features cause multipath issues and block radio signals, hence impacting the ability of a mobile station to estimate its position. Further work should indicate to what extend the terrain influences a position estimate

The impact of the magnetic field on the propagation of radio signals has been ignored. However, it is shown that strong crustal magnetic fields exist, especially in the southern hemisphere. These magnetic fields affect the refraction and group delays of radio signals. Further investigations should be made to determine their ultimate impact on the precision of a position estimate.

The nightside ionosphere of Mars has not been treated in this work. As was shown, the nightside is far more feeble than the dayside ionosphere. Nevertheless, radio signals can still be reflected and refracted, albeit at a significantly lower frequency than for the dayside. This ultimately impacts the design of the beacons, considering larger antennae are required to support a lower frequency.

It is expected that more performance improvements can be obtained by combining some of the localization technologies. It is not unreasonable to assume that a beacon is able to detect more than one property of an incoming radio signal, thus supporting the ability to use more than one property to estimate a position. It is expected that this will further enhance the performance, as well as provide a more stable position estimate. In fact, the combination of localization technologies might solve the aforementioned non-ideal situations which have not been treated in this work.

The beacons have been assumed infallible. In reality, such is not the case. Therefore, it is recommended to study to what extend a performance estimate degrades when one or more beacons degrade.

Then, the antenna design can be optimized to allow for a higher maximum path loss. In this work, only a simple short dipole is assumed. By equipping the beacons with more novel technologies, the maximum allowable path loss might be raised.

Finally, the TDOA technology has been mentioned, but not analyzed in detail. It is expected that this technology is able to mitigate some of the problems of a TOA -approach, like the issue of time synchronization between beacons.

Conclusively, the outlook for a ground-based beacon system is promising. Further research will have to show in more detail the performance of such a system under various extreme cases. Nevertheless, the approach proposed in this research is a worthy alternative to a space-based system, by providing position estimates with the use of ground-based beacons.

Bibliography

Acuña, M. H. (1999). Global distribution of crustal magnetization discovered by the mars global surveyor MAG/ER experiment. *Science*, 284(5415):790–793.

Acuña, M. H., Connerney, J. E. P., Wasilewski, P., Lin, R. P., Mitchell, D., Anderson, K. A., Carlson, C. W., McFadden, J., Rème, H., Mazelle, C., Vignes, D., Bauer, S. J., Cloutier, P., and Ness, N. F. (2001). Magnetic field of mars: Summary of results from the aerobraking and mapping orbits. *J. Geophys. Res.*, 106(E10):23403.

Annis, J. (1987). The atmosphere of mars and optical communications. *TMO Progress Report*, 42-91.

Balanis, C. (1996). *Antenna Theory: Analysis and Design, 2nd Edition*. Wiley.

Barlow, N. (2008). *Mars: An Introduction to its Interior, Surface and Atmosphere (Cambridge Planetary Science)*. Cambridge University Press.

Bell, D., Cesarone, R., Ely, T., Edwards, C., and Townes, S. (2000). Mars network: a mars orbiting communications and navigation satellite constellation.

Bossler, J. D. and Campbell, J. B. (2010). *Manual of Geospatial Science and Technology, Second Edition*. CRC Press.

Bougher, S., Engel, S., Hinson, D., and Forbes, J. (2001). Mars global surveyor radio science electron density profiles : Neutral atmosphere implications. *Geophysical Research Letters*, 28(16):3091–3094.

Carr, J. (2001). *Practical Antenna Handbook*. McGraw-Hill/TAB Electronics.

Chen, R. H., Cravens, T. E., and Nagy, A. F. (1978). The martian ionosphere in light of the viking observations. *J. Geophys. Res.*, 83(A8):3871.

Davies, K. (1969). *Ionospheric Radio Waves*. Blaisdell Publishing Company.

Drake, B. (2009). Mars human exploration mission dra 5.0. Technical report, NASA-SP-2009-566, July.

Duru, F., Gurnett, D. A., Morgan, D. D., Modolo, R., Nagy, A. F., and Najib, D. (2008). Electron densities in the upper ionosphere of mars from the excitation of electron plasma oscillations. *J. Geophys. Res.*, 113(A7).

Edwards, C., Adams, J., Bell, D., Cesarone, R., DePaula, R., Durning, J., Ely, T., Leung, R., McGraw, C., and Rosell, S. (2001). Strategies for telecommunications and navigation in support of mars exploration. *Acta Astronautica*, 48(5-12):661–668.

Forget, F. and Millour, E. (2014). Mars climate database 5.1.

Fox, J. (2004). Advances in the aeronomy of venus and mars. *Advances in Space Research*, 33(2):132–139.

Fry, C. and Yowell, R. (1994). Hf radio on mars. *Commun. Q*, 4(2):13–23.

Gentile, C., Alsindi, N., Raulefs, R., and Teolis, C. (2013). *Geolocation Techniques*. Springer Science + Business Media.

Glassner, A. S. (1989). *An Introduction to Ray Tracing*. Academic Press.

Graf, J., Zurek, R., Eisen, H., Jai, B., Johnston, M., and DePaula, R. (2005). The mars reconnaissance orbiter mission. *Acta Astronautica*, 57(2-8):566–578.

Greeley, R. and Haberle, R. (1991). Sand and dust on mars.

Grinstead, C. M. and Snell, J. L. (2012). *Introduction to Probability: Second Revised Edition*. American Mathematical Society.

Harri, A., Pellinen, R., Uspensky, M., Siili, T., Linkin, V., Lipatov, A., Savijärvi, H., Vorontsov, V., and Ivankov, A. (2006). Metnet-atmospheric science network for mars. *Mars Atmosphere Modelling and Observations, edited by F. Forget et al*, page 724.

Harris, F. (1978). On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, *66*(1):51–83.

Heggy, E., Clifford, S., Younsi, A., Miane, J., Carley, R., and Morris, R. (2007). On the Dielectric Properties of Dust and Ice-Dust Mixtures: Experimental Characterization of the Martian Polar-layered Deposits Analog Materials. 38:1756.

Hess, S. L., Ryan, J. A., Tillman, J. E., Henry, R. M., and Leovy, C. B. (1980). The annual cycle of pressure on mars measured by viking landers 1 and 2. *Geophysical Research Letters*, 7(3):197–200.

Ho, C., Golshan, N., and Kliore, A. (2002). *Radio Wave Propagation Handbook For Communication on and Around Mars*. Jet Propulsion Laboratory, California Institute of Technology.

Ho, C., Sue, M. K., and Golshan, N. (2001). Martian atmospheric effects on radio wave propagation.

ITU-R P.368-7 (1992). Ground-wave propagation curves for frequencies between 10 khz and 30 mhz. *ITU Radiocommunication Assembly*.

ITU-R P.453-6 (1997). The radio refractive index: Its formula and refractivity data. *ITU Radiocommunication Assembly*.

Ivanov, A. B. and Muhleman, D. O. (1998). Opacity of the martian atmosphere from the mars orbiter laser altimeter (MOLA) observations. *Geophysical Research Letters*, 25(24):4417–4420.

Karcz, J., Davis, S., Aftosmis, M., Allen, G., Bakhtian, N., Dyakonov, A., Edquist, K., Glass, B., Gonzales, A., Heldmann, J., Lemke, L., Marinova, M., McKay, C., Stoker, C., Wooster, P., and Zarchi, K. (2012). Red Dragon: Low-Cost Access to the Surface of Mars Using Commercial Capabilities. *LPI Contributions*, 1679:4315.

Kay, S. (1993). *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory (v. 1)*. Prentice Hall.

Kayton, M. and Fried, W. R. (1997). *Avionics Navigation Systems, 2nd Edition*. John Wiley & Sons, Inc.

Kelso, J. (1964). *Radio Ray Propagation in the Ionosphere*. McGraw-Hill.

Kivelson, H. and Russel, C. (1995). *Introduction to Space Physics*. Cambridge University Press.

Laby, K. . (2015). Tables of physical & chemical constants.

Larson, W. and Wertz, J. (1999). *Space Mission Analysis and Design, 3rd edition (Space Technology Library, Vol. 8)*. Microcosm.

Mahajan, K. K., Lodhi, N. K., and Singh, S. (2009). Ionospheric effects of solar flares at mars. *Geophysical Research Letters*, 36(15):n/a–n/a.

Mahajan, V. N. (1998). *Optical Imaging and Aberrations*. SPIE.

Makovsky, A., Ilott, P., and Taylor, J. (2009). Mars science laboratory telecommunications system design. *DESCANSO Design and Performance Summary Series*, Article 14.

Maltagliati, L., Montmessin, F., Fedorova, A., Korablev, O., Forget, F., and Bertaux, J.-L. (2011). Evidence of water vapor in excess of saturation in the atmosphere of mars. *Science*, 333(6051):1868–1871.

Melnik, O. and Parrot, M. (1999). Propagation of electromagnetic waves through the martian ionosphere. *J. Geophys. Res.*, 104(A6):12705.

Mendillo, M. (2006). Effects of solar flares on the ionosphere of mars. *Science*, 311(5764):1135–1138.

Mendillo, M., Pi, X., Smith, S., Martinis, C., Wilson, J., and Hinson, D. (2004). Ionospheric effects upon a satellite navigation system at mars. *Radio Sci.*, 39(2):n/a–n/a.

Millington, G. (1949). Ground-wave propagation over an inhomogeneous smooth earth. *Proceedings of the IEE - Part III: Radio and Communication Engineering*, 96(39):53–64.

Molina-Cuberos, G. J. (2002). Ion-neutral chemistry model of the lower ionosphere of mars. *J. Geophys. Res.*, 107(E5).

Molina-Cuberos, G. J., Witasse, O., Lebreton, J.-P., Rodrigo, R., and López-Moreno, J. J. (2003). Meteoric ions in the atmosphere of mars. *Planetary and Space Science*, 51(3):239–249.

Morgan, D. D., Gurnett, D. A., Kirchner, D. L., Fox, J. L., Nielsen, E., and Plaut, J. J. (2008). Variation of the martian ionospheric electron density from mars express radar soundings. *J. Geophys. Res.*, 113(A9).

Němec, F., Morgan, D. D., Gurnett, D. A., and Duru, F. (2010). Nightside ionosphere of mars: Radar soundings by the mars express spacecraft. *J. Geophys. Res.*, 115(E12).

Nielsen, E., Morgan, D., Kirchner, D., Plaut, J., and Picardi, G. (2007). Absorption and reflection of radio waves in the martian ionosphere. *Planetary and Space Science*, 55(7-8):864–870.

Parker, T. (2013). Localization and 'contextualization' of curiosity in gale crater, and other landed mars missions.

Pesnell, W. D. and Grebowsky, J. (2000). Meteoric magnesium ions in the martian atmosphere. *J. Geophys. Res.*, 105(E1):1695.

Pollack, J. B., Colburn, D. S., Flasar, F. M., Kahn, R., Carlston, C. E., and Pidek, D. (1979). Properties and effects of dust particles suspended in the martian atmosphere. *J. Geophys. Res.*, 84(B6):2929.

Powell, M., Crockett, T., Fox, J., Joswig, J., Norris, J., and Rabe, K. (2006). Targeting and localization for mars rover operations. In *2006 IEEE International Conference on Information Reuse & Integration*. IEEE.

Qi, Y. and Kobayashi, H. (2002). Cramer-rao lower bound for geolocation in non-line-of-sight environment. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 3, pages III–2473. IEEE.

Riedl, M. J. (2009). *Optical Design*. SPIE.

Roberts, J. H., Lillis, R. J., and Manga, M. (2009). Giant impacts on early mars and the cessation of the martian dynamo. *J. Geophys. Res.*, 114(E4).

Safaeinili, A., Kofman, W., Mouginot, J., Gim, Y., Herique, A., Ivanov, A. B., Plaut, J. J., and Picardi, G. (2007). Estimation of the total electron content of the martian ionosphere using radar sounder surface echoes. *Geophysical Research Letters*, 34(23):n/a–n/a.

Safaeinili, A., Kofman, W., Nouvel, J.-F., Herique, A., and Jordan, R. (2003). Impact of mars ionosphere on orbital radar sounder operation and data processing. *Planetary and Space Science*, 51(7-8):505–515.

Schofield, J. (1997). The mars pathfinder atmospheric structure investigation/meteorology (ASI/MET) experiment. *Science*, 278(5344):1752–1758.

Schunk, R. and Nagy, A. (2009). *Ionospheres: Physics, Plasma Physics, and Chemistry (Cambridge Atmospheric and Space Science Series)*. Cambridge University Press.

Shirley, J. and Fairbridge, R. (1997). *Encyclopedia of Planetary Sciences*. Chapman & Hall.

Smith, E. and Flock, W. (1986). Propagation through martian dust at 8.5 and 32 ghz. *TDA Progress Report*, 42-87.

Smith, M. D. (2009). THEMIS observations of mars aerosol optical depth from 2002–2008. *Icarus*, 202(2):444–452.

Volpe, R. (2014). 2014 robotics activities at jpl.

Wang, J. and Nielsen, E. (2003). Behavior of the martian dayside electron density peak during global dust storms. *Planetary and Space Science*, 51(4-5):329–338.

Whitted, T. (1980). An improved illumination model for shaded display.

Williams, D. D. R. (2014 (accessed 26 March 2015)). *Mars Fact Sheet*. `http://nssdc.gsfc.nasa.gov/planetary/factsheet/marsfact.html`.

Witasse, O., Nouvel, J.-F., Lebreton, J.-P., and Kofman, W. (2001). HF radio wave attenuation due to a meteoric layer in the atmosphere of mars. *Geophysical Research Letters*, 28(15):3039–3042.

Withers, P. (2009). A review of observed variability in the dayside ionosphere of mars. *Advances in Space Research*, 44(3):277–307.

Withers, P. (2011). Attenuation of radio signals by the ionosphere of mars: Theoretical development and application to MARSIS observations. *Radio Sci.*, 46(2):n/a–n/a.

Withers, P., Fallows, K., and Matta, M. (2014). Predictions of electron temperatures in the mars ionosphere and their effects on electron densities. *Geophysical Research Letters*, 41(8):2681–2686.

Withers, P., Fillingim, M. O., Lillis, R. J., Häusler, B., Hinson, D. P., Tyler, G. L., Pätzold, M., Peter, K., Tellmann, S., and Witasse, O. (2012). Observations of the nightside ionosphere of mars by the mars express radio science experiment (MaRS). *J. Geophys. Res.*, 117(A12):n/a–n/a.

Woan, G. (2000). *The Cambridge Handbook of Physics Formulas*. Cambridge University Press (CUP).

Zent, A. P., Hecht, M. H., Cobos, D. R., Wood, S. E., Hudson, T. L., Milkovich, S. M., DeFlores, L. P., and Mellon, M. T. (2010). Initial results from the thermal and electrical conductivity probe (TECP) on phoenix. *J. Geophys. Res.*, 115.

Zubrin, R. (1992). Long range mobility on mars. *Journal of the British Interplanetary Society*, 45:203–210.

Zyren, J. and Petrick, A. (1998). Tutorial on basic link budget analyis. *Intersil. Application Note AN9804.1*.

Index

A — Derivation of the Jacobian

A crucial step in the determination of the accuracy of a localization system is to derivate the formula for the measurement matrix. Both the beacon and the mobile station have a certain position β_{BC}, λ_{BC} and β_{MS}, λ_{MS} respectively (see figure 5.4). In order to obtain the Jacobian of the measurement matrix, it needs to be derived with respect to the parameter vector $\boldsymbol{\theta}$. Herein, the following values are relevant:

- Orthodromic distance to a beacon n: d_n
- Parameter vector: $\boldsymbol{\theta} = [\beta\ \lambda]$
- Longitudinal separation between beacon and mobile station: $\Delta\lambda = |\lambda_{BC} - \lambda_{MS}|$ (refer to figure 5.2.1)

A.1 Derivation for the TOA approach

The measured value for a TOA approach is the propagation time, τ. The Jacobian of this value follows from the derivation of the propagation time from the parameter vector. The derivation starts with the relationship between propagation time and orthodromic distance. It has been found that this relationship is a second-order polynomial, with constants a, b and c:

$$\tau_n = a d_n^3 + b d_n^2 + c d_n + d \tag{A.1}$$

It is known that the orthodromic distance depends on the angular separation and planetary radius by $d_n = \delta\phi \cdot R_M$. Furthermore, the angular separation is defined as:

$$\delta\phi = \cos^{-1}\left(\sin(\beta_B)\sin(\beta_M) + \cos(\beta_B)\cos(\beta_M)\cos(\Delta\lambda)\right) \tag{A.2}$$

Combining equations A.1 and A.2 yields:

$$\tau_n = a\left(R_M \cos^{-1}(x)\right)^3 + b\left(R_M \cos^{-1}(x)\right)^2 + cR_M \cos^{-1}(x) + d \tag{A.3}$$

$$x = \sin(\beta_B)\sin(\beta_M) + \cos(\beta_B)\cos(\beta_M)\cos(\Delta\lambda) \tag{A.4}$$

A.1.1 Derivation with respect to the latitude

The derivation of τ_n with respect to the latitude, $\partial\tau/\partial\beta$, starts with equation A.1. The derivation is done with integration by parts:

$$\frac{\partial\tau}{\partial\beta} = \left(3ad_n^2 + 2bd_n + c\right)\frac{\partial d_n}{\partial\beta} \tag{A.5}$$

$$\frac{\partial d_n}{\partial\beta} = \frac{\partial}{\partial\beta} R_M \cos^{-1}(x) = \frac{-R_M}{\sqrt{1-x^2}}\frac{\partial x}{\partial\beta} \tag{A.6}$$

$$\begin{aligned}\frac{\partial x}{\partial\beta} &= \frac{\partial}{\partial\beta}\left(\cos(\beta_{BC})\cos(\beta_{MS})\cos(\Delta\lambda) + \sin(\beta_{BC})\sin(\beta_{MS})\right) \\ &= \left(\cos(\beta_{BC})\sin(\beta_{MS})\cos(\Delta\lambda) + \sin(\beta_M S)\cos(\beta_{BC})\right)\end{aligned} \tag{A.7}$$

Inserting the partial derivatives from equations A.6 and A.7 into equation A.5 yields:

$$\begin{aligned}\frac{\partial\tau}{\partial\beta} &= \left(3ad_n^2+2bd_n+c\right)\frac{-R_M}{\sqrt{1-x^2}} \\ &\quad\cdot\left(\cos(\beta_{BC})\sin(\beta_{MS})\cos(\Delta\lambda)+\sin(\beta_M S)\cos(\beta_{BC})\right)\end{aligned} \tag{A.8}$$

A.1.2 Derivation with respect to the longitude

The derivation of τ_n with respect to the longitude, $\partial\tau/\partial\lambda$, is largely comparable to the derivation with respect to the latitude. However, due to the composition of the angular separation, the derivation itself is less complex. Note, that the sign changes, depending on the values inside $\Delta\lambda$.
Let:

$$\frac{\partial\tau}{\partial\lambda} = \left(3ad_n^2+2bd_n+c\right)\frac{\partial d_n}{\partial\lambda} \tag{A.9}$$

Then:

$$\frac{\partial d_n}{\partial\lambda} = \frac{\partial}{\partial\lambda}R_M\cos^{-1}(x) = \frac{-R_M}{\sqrt{1-x^2}}\frac{\partial x}{\partial\lambda} \tag{A.10}$$

$$\begin{aligned}\frac{\partial x}{\partial\lambda} &= \frac{\partial}{\partial\lambda}\left(\cos(\beta_{BC})\cos(\beta_{MS})\cos(\Delta\lambda)+\sin(\beta_{BC})\sin(\beta_{MS})\right) \\ &= -\cos(\beta_{BC})\cos(\beta_{MS})\sin(\Delta\lambda)\end{aligned} \tag{A.11}$$

(A.12)

Again, inserting equations A.10 and A.11 into A.9 yields:

$$\frac{\partial\tau}{\partial\lambda} = \left(3ad_n^2+2bd_n+c\right)\frac{R_M}{\sqrt{1-x^2}}\cos(\beta_{BC})\cos(\beta_{MS})\sin(\Delta\lambda) \tag{A.13}$$

Equation A.13 changes sign when $\lambda_{BC} < \lambda_{MS}$ due to the derivation of $\Delta\lambda$. Then:

$$\frac{\partial\tau}{\partial\lambda} = -\left(3ad_n^2+2bd_n+c\right)\frac{R_M}{\sqrt{1-x^2}}\cos(\beta_{BC})\cos(\beta_{MS})\sin(\Delta\lambda) \tag{A.14}$$

A.2 Derivation for the RSS approach

For the RSS approach, the signal strength P_r is measured. As with the TOA approach, the Jacobian is found by derivating P_r with respect to the parameter vector. The relationship between the measurement value and the distance is a second-order exponential function:

$$P_r = a_1\exp\left(-\left(\frac{d_n+b_1}{c_1}\right)^2\right)+a_2\exp\left(-\left(\frac{d_n+b_2}{c_2}\right)^2\right) \tag{A.15}$$

A.2.1 Derivation with respect to the latitude

The derivation with respect to the latitude is found by derivating equation A.15 with respect to β:

$$\begin{aligned}\frac{\partial P_R}{\partial\beta} &= \frac{2a_1}{c_1^2}\left(d_n-b_1\right)\exp\left(-\left(\frac{d_n+b_1}{c_1}\right)^2\right)\frac{\partial d_n}{\partial\beta} \\ &+ \frac{2a_2}{c_2^2}\left(d_n-b_2\right)\exp\left(-\left(\frac{d_n+b_2}{c_2}\right)^2\right)\frac{\partial d_n}{\partial\beta}\end{aligned} \tag{A.16}$$

The partial derivation of d_n w.r.t. β is taken from equation A.6. Inserting this equation into equation A.16 gives:

$$\frac{\partial P_R}{\partial \beta} = \left[\frac{2a_1}{c_1^2}(d_n - b_1) \exp\left(-\left(\frac{d_n + b_1}{c_1} \right)^2 \right) + \frac{2a_2}{c_2^2}(d_n - b_2) \exp\left(-\left(\frac{d_n + b_2}{c_2} \right)^2 \right) \right] \cdot \frac{-R_M}{\sqrt{1-x^2}} \left(\cos(\beta_{BC}) \sin(\beta_{MS}) \cos(\Delta\lambda) + \sin(\beta_M S) \cos(\beta_{BC}) \right) \tag{A.17}$$

A.2.2 Derivation with respect to the longitude

The derivation of the received signal strength w.r.t. the longitude follows the same process as the derivation w.r.t. the latitude. Except, in this case, equation A.16 is rewritten as:

$$\begin{aligned} \frac{\partial P_R}{\partial \lambda} &= \frac{2a_1}{c_1^2}(d_n - b_1) \exp\left(-\left(\frac{d_n + b_1}{c_1} \right)^2 \right) \frac{\partial d_n}{\partial \lambda} \\ &+ \frac{2a_2}{c_2^2}(d_n - b_2) \exp\left(-\left(\frac{d_n + b_2}{c_2} \right)^2 \right) \frac{\partial d_n}{\partial \lambda} \end{aligned} \tag{A.18}$$

Inserting equation A.10 into equation A.18 and rewriting gives:

$$\frac{\partial P_R}{\partial \lambda} = \left[\frac{2a_1}{c_1^2}(d_n - b_1) \exp\left(-\left(\frac{d_n + b_1}{c_1} \right)^2 \right) + \frac{2a_2}{c_2^2}(d_n - b_2) \exp\left(-\left(\frac{d_n + b_2}{c_2} \right)^2 \right) \right] \cdot \frac{R_M}{\sqrt{1-x^2}} \cos(\beta_{BC}) \cos(\beta_{MS}) \sin(\Delta\lambda) \tag{A.19}$$

Equation A.19 changes sign for $\lambda_{BC} < \lambda_{MS}$:

$$\frac{\partial P_R}{\partial \lambda} = -\left[\frac{2a_1}{c_1^2}(d_n - b_1) \exp\left(-\left(\frac{d_n + b_1}{c_1} \right)^2 \right) + \frac{2a_2}{c_2^2}(d_n - b_2) \exp\left(-\left(\frac{d_n + b_2}{c_2} \right)^2 \right) \right] \cdot \frac{R_M}{\sqrt{1-x^2}} \cos(\beta_{BC}) \cos(\beta_{MS}) \sin(\Delta\lambda) \tag{A.20}$$

A.3 Derivation for the AoA approach

The derivation for the AoA approach follows the same procedure as the procedure for the other two methods described above. It is found that the model fit for an AoA approach has the following definition:

$$\gamma_a = a \exp(bd) + c \exp(gd) \tag{A.21}$$

A.3.1 Derivation with respect to the latitude

The derivation of equation A.21 with respect to the latitude yields:

$$\frac{\partial \gamma}{\partial \beta} = \left[ab \exp^{bd_n} + cg \exp^{gd_n} \right] \frac{\partial d_n}{\partial \beta} \tag{A.22}$$

Using the same procedure as before, equation A.22 is combined with equation A.6, which produces:

$$\begin{aligned}\frac{\partial\gamma}{\partial\beta} &= \left[ab\exp^{bd_n} + cg\exp^{gd_n}\right]\frac{-R_M}{\sqrt{1-x^2}} \\ &\cdot(\cos(\beta_{BC})\sin(\beta_{MS})\cos(\Delta\lambda) + \sin(\beta_{MS})\cos(\beta_{BC}))\end{aligned} \tag{A.23}$$

A.3.2 Derivation with respect to the longitude

The derivation with respect to the longitude follows again from equation A.21:

$$\frac{\partial\gamma}{\partial\lambda} = \left[ab\exp^{bd_n} + cg\exp^{gd_n}\right]\frac{\partial d_n}{\partial\lambda} \tag{A.24}$$

Combining equations A.24 and A.10 yields:

$$\begin{aligned}\frac{\partial\gamma}{\partial\lambda} &= \left[ab\exp^{bd_n} + cg\exp^{gd_n}\right]\frac{R_M}{\sqrt{1-x^2}} \\ &\cos(\beta_{BC})\cos(\beta_{MS})\sin(\Delta\lambda)\end{aligned} \tag{A.25}$$

For the case where $\lambda_{BC} < \lambda_{MS}$:

$$\begin{aligned}\frac{\partial\gamma}{\partial\lambda} &= -\left[ab\exp^{bd_n} + cg\exp^{gd_n}\right]\frac{R_M}{\sqrt{1-x^2}} \\ &\cos(\beta_{BC})\cos(\beta_{MS})\sin(\Delta\lambda)\end{aligned} \tag{A.26}$$

B — Measure of Precision values

Table B.1: Values of the measure of precision M_P, as found by the simulation. All values for all technologies and frequencies, for the nominal, max *and* min *scenario are listed.*

Scenario	Approach	Frequency	M_P
Nominal	ToA	$2.5MHz$	$M_P = diag(2.22 \cdot 10^9, 2.22 \cdot 10^9, 2.22 \cdot 10^9, 2.22 \cdot 10^9)$
		$3.5MHz$	$M_P = diag(1.69 \cdot 10^9, 1.69 \cdot 10^9, 1.69 \cdot 10^9, 1.69 \cdot 10^9)$
		$4.5MHz$	$M_P = diag(6.94 \cdot 10^8, 6.94 \cdot 10^8, 6.94 \cdot 10^8, 6.94 \cdot 10^8)$
		$5.5MHz$	$M_P = diag(3.35 \cdot 10^9, 3.35 \cdot 10^9, 3.35 \cdot 10^9, 3.35 \cdot 10^9)$
Nominal	AoA	$2.5MHz$	$M_P = diag(1.49 \cdot 10^4, 1.49 \cdot 10^4, 1.49 \cdot 10^3, 1.49 \cdot 10^3)$
		$3.5MHz$	$M_P = diag(4.68 \cdot 10^4, 4.68 \cdot 10^4, 4.68 \cdot 10^3, 4.68 \cdot 10^3)$
		$4.5MHz$	$M_P = diag(6.57 \cdot 10^4, 6.57 \cdot 10^4, 6.57 \cdot 10^3, 6.57 \cdot 10^3)$
		$5.5MHz$	$M_P = diag(1.68 \cdot 10^4, 1.68 \cdot 10^4, 1.68 \cdot 10^3, 1.68 \cdot 10^3)$
Nominal	RSS	$2.5MHz$	$M_P = diag(0.52, 0.52, 0.52, 0.52)$
		$3.5MHz$	$M_P = diag(2.66, 2.66, 2.66, 2.66)$
		$4.5MHz$	$M_P = diag(0.70, 0.70, 0.70, 0.70)$
		$5.5MHz$	$M_P = diag(0.11, 0.11, 0.11, 0.11)$
Max	ToA	$2.5MHz$	$M_P = diag(2.41 \cdot 10^9, 2.41 \cdot 10^9, 2.41 \cdot 10^9, 2.41 \cdot 10^9)$
		$3.5MHz$	$M_P = diag(1.10 \cdot 10^9, 1.10 \cdot 10^9, 1.10 \cdot 10^9, 1.10 \cdot 10^9)$
		$4.5MHz$	$M_P = diag(2.14 \cdot 10^9, 2.14 \cdot 10^9, 2.14 \cdot 10^9, 2.14 \cdot 10^9)$
		$5.5MHz$	$M_P = diag(1.65 \cdot 10^9, 1.65 \cdot 10^9, 1.65 \cdot 10^9, 1.65 \cdot 10^9)$
Max	AoA	$2.5MHz$	$M_P = diag(1.78 \cdot 10^4, 1.78 \cdot 10^4, 1.78 \cdot 10^4, 1.78 \cdot 10^4)$
		$3.5MHz$	$M_P = diag(7.81 \cdot 10^2, 7.81 \cdot 10^2, 7.81 \cdot 10^2, 7.81 \cdot 10^2)$
		$4.5MHz$	$M_P = diag(3.72 \cdot 10^3, 3.72 \cdot 10^3, 3.72 \cdot 10^3, 3.72 \cdot 10^3)$
		$5.5MHz$	$M_P = diag(9.59 \cdot 10^2, 9.59 \cdot 10^2, 9.59 \cdot 10^2, 9.59 \cdot 10^2)$
Max	RSS	$2.5MHz$	$M_P = diag(9.25 \cdot 10^{-1}, 9.25 \cdot 10^{-1}, 9.25 \cdot 10^{-1}, 9.25 \cdot 10^{-1})$
		$3.5MHz$	$M_P = diag(1.62 \cdot 10^{-1}, 1.62 \cdot 10^{-1}, 1.62 \cdot 10^{-1}, 1.62 \cdot 10^{-1})$
		$4.5MHz$	$M_P = diag(1.90 \cdot 10^{-1}, 1.90 \cdot 10^{-1}, 1.90 \cdot 10^{-1}, 1.90 \cdot 10^{-1})$
		$5.5MHz$	$M_P = diag(2.32 \cdot 10^{-1}, 2.32 \cdot 10^{-1}, 2.32 \cdot 10^{-1}, 2.32 \cdot 10^{-1})$
Min	ToA	$2.5MHz$	$M_P = diag(3.83 \cdot 10^9, 3.83 \cdot 10^9, 3.83 \cdot 10^9, 3.83 \cdot 10^9)$
		$3.5MHz$	$M_P = diag(2.35 \cdot 10^9, 2.35 \cdot 10^9, 2.35 \cdot 10^9, 2.35 \cdot 10^9)$
		$4.5MHz$	$M_P = diag(1.90 \cdot 10^9, 1.90 \cdot 10^9, 1.90 \cdot 10^9, 1.90 \cdot 10^9)$
		$5.5MHz$	$M_P = diag(1.62 \cdot 10^{10}, 1.62 \cdot 10^{10}, 1.62 \cdot 10^{10}, 1.62 \cdot 10^{10})$
Min	AoA	$2.5MHz$	$M_P = diag(2.28 \cdot 10^4, 2.28 \cdot 10^4, 2.28 \cdot 10^4, 2.28 \cdot 10^4)$
		$3.5MHz$	$M_P = diag(8.25 \cdot 10^4, 8.25 \cdot 10^4, 8.25 \cdot 10^4, 8.25 \cdot 10^4)$
		$4.5MHz$	$M_P = diag(3.31 \cdot 10^4, 3.31 \cdot 10^4, 3.31 \cdot 10^4, 3.31 \cdot 10^4)$
		$5.5MHz$	$M_P = diag(1.10 \cdot 10^4, 1.10 \cdot 10^4, 1.10 \cdot 10^4, 1.10 \cdot 10^4)$
Min	RSS	$2.5MHz$	$M_P = diag(1.77 \cdot 10^{-1}, 1.77 \cdot 10^{-1}, 1.77 \cdot 10^{-1}, 1.77 \cdot 10^{-1})$
		$3.5MHz$	$M_P = diag(2.86 \cdot 10^{-1}, 2.86 \cdot 10^{-1}, 2.86 \cdot 10^{-1}, 2.86 \cdot 10^{-1})$
		$4.5MHz$	$M_P = diag(0.10 \cdot 10^{-1}, 0.10 \cdot 10^{-1}, 0.10 \cdot 10^{-1}, 0.10 \cdot 10^{-1})$
		$5.5MHz$	$M_P = diag(0.69 \cdot 10^{-1}, 0.69 \cdot 10^{-1}, 0.69 \cdot 10^{-1}, 0.69 \cdot 10^{-1})$

Table B.2: Values of the measure of precision $\boldsymbol{M_P}$, as found by the simulation. All values for all technologies and frequencies, for the dust storm *and* sep *are listed.*

Scenario	Approach	Frequency	$\boldsymbol{M_P}$
Duststorm	ToA	$2.5MHz$	$M_P = diag(1.96 \cdot 10^9, 1.96 \cdot 10^9, 1.96 \cdot 10^9, 1.96 \cdot 10^9)$
		$3.5MHz$	$M_P = diag(1.16 \cdot 10^9, 1.16 \cdot 10^9, 1.16 \cdot 10^9, 1.16 \cdot 10^9)$
		$4.5MHz$	$M_P = diag(0.85 \cdot 10^9, 0.85 \cdot 10^9, 0.85 \cdot 10^9, 0.85 \cdot 10^9)$
		$5.5MHz$	$M_P = diag(7.23 \cdot 10^9, 7.23 \cdot 10^9, 7.23 \cdot 10^9, 7.23 \cdot 10^9)$
Duststorm	AoA	$2.5MHz$	$M_P = diag(1.58 \cdot 10^4, 1.58 \cdot 10^4, 1.58 \cdot 10^4, 1.58 \cdot 10^4)$
		$3.5MHz$	$M_P = diag(1.45 \cdot 10^4, 1.45 \cdot 10^4, 1.45 \cdot 10^4, 1.45 \cdot 10^4)$
		$4.5MHz$	$M_P = diag(2.11 \cdot 10^4, 2.11 \cdot 10^4, 2.11 \cdot 10^4, 2.11 \cdot 10^4)$
		$5.5MHz$	$M_P = diag(0.72 \cdot 10^4, 0.72 \cdot 10^4, 0.72 \cdot 10^4, 0.72 \cdot 10^4)$
Duststorm	RSS	$2.5MHz$	$M_P = diag(2.80, 2.80, 2.80, 2.80)$
		$3.5MHz$	$M_P = diag(7.94, 7.94, 7.94, 7.94)$
		$4.5MHz$	$M_P = diag(11.1, 11.1, 11.1, 11.1)$
		$5.5MHz$	$M_P = diag(2.74, 2.74, 2.74, 2.74)$
SEP	ToA	$2.5MHz$	$M_P = diag(2.83 \cdot 10^9, 2.83 \cdot 10^9, 2.83 \cdot 10^9)$
		$3.5MHz$	$M_P = diag(0.49 \cdot 10^9, 0.49 \cdot 10^9, 0.49 \cdot 10^9)$
		$4.5MHz$	$M_P = diag(0.53 \cdot 10^9, 0.53 \cdot 10^9, 0.53 \cdot 10^9)$
SEP	AoA	$2.5MHz$	$M_P = diag(2.16 \cdot 10^4, 2.16 \cdot 10^4, 2.16 \cdot 10^4)$
		$3.5MHz$	$M_P = diag(2.29 \cdot 10^2, 2.29 \cdot 10^2, 2.29 \cdot 10^2)$
		$4.5MHz$	$M_P = diag(2.47 \cdot 10^2, 2.47 \cdot 10^2, 2.47 \cdot 10^2)$
SEP	RSS	$2.5MHz$	$M_P = diag(3.51 \cdot 10^{-1}, 3.51 \cdot 10^{-1}, 3.51 \cdot 10^{-1})$
		$3.5MHz$	$M_P = diag(4.96 \cdot 10^{-4}, 4.96 \cdot 10^{-4}, 4.96 \cdot 10^{-4})$
		$4.5MHz$	$M_P = diag(6.77 \cdot 10^{-4}, 6.77 \cdot 10^{-4}, 6.77 \cdot 10^{-4})$

C — Manual of the Simulation

C.1 Installation

This section treats how to prepare your system so that the simulation can be installed and executed. Note that all steps are necessary, even if you downloaded the binary. The simulation tool makes use of external libraries, which need to be available on your system in order for the simulation to work. At the time of writing, only Linux is supported.

The source code of the simulation is available as a GitHub repository, under the MIT license. The code can be found on:
`http://github.com/rvangijlswijk/ionospheric-ray-tracer`

C.1.1 Prerequisites

This project makes use of the following packages:

- GTest: `http://code.google.com/p/googletest/`
- Boost libraries: `http://www.boost.org/`
- scons: `http://www.scons.org/`
- jsoncpp: `https://github.com/open-source-parsers/jsoncpp`

GTest is included in the source, but it is required to have both the boost and scons json libraries installed locally.

C.1.2 Installation on Linux

The following section treats how to prepare your system for this package if you want to compile it from source. This procedure is tested on systems running *Ubuntu 14.04* and *Ubuntu 14.10*.

Step 1 Install git:

```
sudo apt-get install git
```

Step 2 Ensure that the required build packages and libraries are installed. The minimum required gcc is gcc 4.9.

```
sudo add-apt-repository ppa:ubuntu-toolchain-r/test
sudo apt-get update
sudo apt-get upgrade
sudo apt-get install build-essential g++ gcc python-dev
    autotools-dev libicu-dev build-essential libbz2-dev
    libstdc++
```

Step 3 Install the boost libraries. The minimum required version is libboost-1.55.

```
sudo aptitude install libboost-all-dev
```

Step 4 The next step is to download and compile the json library, *jsoncpp*. This library requires the software configuration tool *scons*, so that needs to be installed as well. Note that the version number of gcc might vary on your system. In the example below, gcc-4.9.2 is used:

```
sudo apt-get install scons
git clone https://github.com/open-source-parsers/jsoncpp.git
cd jsoncpp
scons platform=linux-gcc
sudo cp libs/linux-gcc-4.9.2/libjson_linux-gcc-4.9.2_libmt.so
    /usr/lib
```

Step 5 The final step is to download the simulator itself. The release versions of the simulator are hosted on the following webpage:
`http://github.com/rvangijlswijk/ionospheric-ray-tracer/releases`
One should download the file `irt.zip` in order to obtain the executable. This zipfile contains the executable, and a set of configuration files, which are required to configure the simulation. The zipfile must be unzipped to some folder on your system. Within this folder, the simulator can be executed.

C.2 Usage

The simulation can be configured in two ways. First, runtime options allow to set certain parameters at the command line. Second, configuration files store the configuration of both the scenario used and the application settings.

This section assumes that you are running a linux machine. The commands will not work on Windows.

C.2.1 Quickstart

The simulation can be executed with the following command:

```
./irt config/scenario_nominal.json
```

Note that this command should be executed in the folder where the binary is downloaded. The command runs the simulation for the nominal scenario. The last parameter always indicates the configuration file for the planetary environment which should be used. In this example, the environment for the *nominal scenario* is used. Furthermore, the simulation assumes that a file exists at the relative path config/config.json. This file is automatically loaded.

The results of the simulation are stored in a csv data file, which can be easily loaded by Matlab or some csv reader. The default location of the data file is *Debug/data.dat*. Ensure that the *Debug* folder exists before running the simulator.

C.2.2 Runtime options

Several options can be passed to the simulation to alter application settings. The options can be provided in the command line as follows:

```
./irt [-opts] scenarioConfig
```

The options are supplied as the first arguments. The scenario configuration file is always the last argument. This configuration file is also obligatory. Table C.1 lists the options that can be supplied.

Table C.1: List of command-line options

Option	Explanation
`-c | --config [path]`	Provide a custom application configuration file, located at `path`
`-i | --iterations [num]`	The number, `num` of consecutive times every ray should be simulated. This is used for the Monte-Carlo approach. An increase in iterations will increase the number of measurements.
`-h | --help`	Print a help text
`-o | --output [path]`	Store the output at the location indicated by `path`. It is advised to have `path` end with `.dat`, i.e. `/path/to/data.dat`
`-p | --parallelism`	The number of threads to use for the simulation
`-v | --verbose`	Verbose, display log output
`-vv`	Very verbose, display both log output and debug output

For example, to run the simulator on four threads, write the output to `mydata.dat` and use the configuration for the *max* scenario, one must use the following command:

```
./irt -o mydata.dat -p 4 config/scenario_max.json
```

C.2.3 Configuration files

The configuration of the simulation and the scenarios are stored in an application and scenario configuration file, respectively. The release download includes a folder, containing one application configuration file and scenario configuration files for several scenarios. These configurations are defined in *json* format. Some of the application configuration parameters can be overridden with the use of runtime options.

An example configuration file is given in codeblock C.1. The parameters that define the planet, its atmosphere and its ionosphere are dictated in this configuration. Parameters are defined in key-value pairs. The key dictates the name of the parameter, i.e. “radius”. The value can be given as number, scientific notation, array or text string. The simulation assumes a certain format for each value, and therefore the format as shown in the example

must be used. For example, the "radius" parameter uses an integer number value. The "surfaceNCO2" parameter uses scientific notation. Note that the latter value is stated in brackets. It is assumed that all parameter names, given in the example, are self-explanatory.

A multilayer ionosphere is defined by adding an extra block with parameters to the "layer parameter", which itself is an array. Example C.1 contains an ionosphere with two layers: the M2 and M1 layer. Further information about the json format is found at `http://json.org/`.

```
{
    "name": "Mars",
    "radius": 3390000,
    "surfaceNCO2": "2.8e17",
    "surfaceTemperature": 200,
    "atmosphere": {
        "start": 1000,
        "end": 40000,
    },
    "ionosphere": {
        "start": 70000,
        "step": 125,
        "end": 250000,
        "layers": [{
            "name" : "M2␣Layer",
            "peakProductionAltitude": 130000,
            "electronPeakDensity": "1.8e11",
            "neutralScaleHeight": 11000,
            "type": "chapman"
        },{
            "name" : "M1␣Layer",
            "peakProductionAltitude": 110000,
            "electronPeakDensity": "9.0e10",
            "neutralScaleHeight": 7600,
            "type": "chapman"
        }]
    }
}
```

Codeblock C.1: An example of a scenario configuration file

C.2.4 Executing from source

The simulation can be executed from source. The main application file is located under `src/core/main.cpp`, as seen from the root of the source code directory structure. The `main.cpp` file is a bootstrapper for the application file in `Application.cpp`. This bootstrapper is configured in both *debug* and *release* mode. In debug mode, all the unittests are run. The simulation itself is run in release mode.